EFFECT OF NON-LINEAR GAIN ON THE RESPONSE OF A POSITION SERVOMECHANISM

R. L. BOWERS

1953

M-31

EFFECT OF NON-LINEAR GAIN ON THE RESPONSE OF A POSITION SERVOMECHANISM

-

R. L. Bowers

EFFECT OF NON-LINEAR GAIN ON THE RESPONSE OF A POSITION SERVOMECHANISM

by

RICHARD LEE BOWERS,
Lieutenant, United States Navy

Submitted in partial fulfillment
of the requirements
for the degree of
MASTER OF SCIENCE
IN
ELECTRICAL ENGINEERING

UNITED STATES NAVAL POSTGRADUATE SCHOOL
Monterey, California
1953

This work is accepted as fulfilling

the thesis requirements for the degree of

MASTER OF SCIENCE

IN

ELECTRICAL ENGINEERING

from the
United States Naval Postgraduate School

PREFACE

The work on this thesis was performed at the United States Naval Postgraduate School at Monterey, California during the period November 1952 to May 1953.

The topic was suggested by Professor George J. Thaler, of the Electrical Engineering Department of the Postgraduate School.

The author wishes to acknowledge the invaluable aid extended, at all times during the preparation of this thesis, by Professor Thaler.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

TABLE OF SYMBOLS AND ABBREVIATIONS

A.C.	Alternating current
D.C.	Direct current
E	Electric potential
f	Frequency, cycles per second
F	Coefficient of viscous damping
j	$\sqrt{-1}$
J	Polar moment of inertia
k	One thousand ohms
K	Gain
M	One million ohms
r_1, r_2	Roots of equation
R	Resistance
t	Time
α	$\frac{F}{2J}$
β	$\sqrt{K/J}$
ϵ	Base of natural logarithms
ε	Error, $\theta_i - \theta_o$
θ	Angular displacement, radians
θ_i	System input
θ_o	System output
ω	Angular velocity, radians per second
τ	Time constant
μfd	Microfarads

SUMMARY

Objective:

To study the effect of a non-linear gain characteristic on the performance of an otherwise linear position servomechanism.

General Methods:

A physical system was constructed, but due to difficulties encountered caused by backlash and static friction, this system could not be used to obtain conclusive results. The analog computer was then used to simulate the system. Various types of gain vs error characteristics were then tested to determine their effect on transient performance.

Conclusions:

It has been proven, theoretically and experimentally, that a gain characteristic which gives increased gain for increased error is of no practical value. The investigation indicates that other types of non-linear gain variation may lead to a system which has improved performance over a linear system.

THE EFFECT OF GAIN ON THE TRANSIENT RESPONSE OF A LINEAR POSITION SERVOMECHANISM

Before going into the behavior of a system using non-linear gain, a general discussion of the behavior of a linear system is in order. For simplicity, a second order linear system will be used to illustrate the effect of gain on the transient response of the system. A block diagram for this type of system is shown in Figure I. The differential equation for this system is:

$$J \frac{d^2\theta_o}{dt^2} + F \frac{d\theta_o}{dt} + K\theta_o = K\theta_i \qquad (1)$$

The characteristic equation has the roots:

$$r_1, r_2 = -\frac{F}{2J} \pm \sqrt{\left(\frac{F}{2J}\right)^2 - \frac{K}{J}} \qquad (2)$$

If the input to the system, θ_i, is a step input of magnitude Θ, and the system is initially at rest with θ_o equal to zero at t equal to zero, the system output will have the following solutions:

for $\frac{K}{J} < \left(\frac{F}{2J}\right)^2$

$$\theta_o = \Theta\left[1 - \epsilon^{-\frac{F}{2J}t}\left(\frac{\frac{F}{2J}}{\sqrt{\left(\frac{F}{2J}\right)^2 - \frac{K}{J}}} \sinh \sqrt{\left(\frac{F}{2J}\right)^2 - \frac{K}{J}}\, t + \cosh \sqrt{\left(\frac{F}{2J}\right)^2 - \frac{K}{J}}\, t\right)\right] \qquad (3)$$

for $\frac{K}{J} = \left(\frac{F}{2J}\right)^2$

$$\theta_o = \Theta\left[1 - \epsilon^{-\frac{F}{2J}t}\left(1 + \frac{F}{2J}t\right)\right] \qquad (4)$$

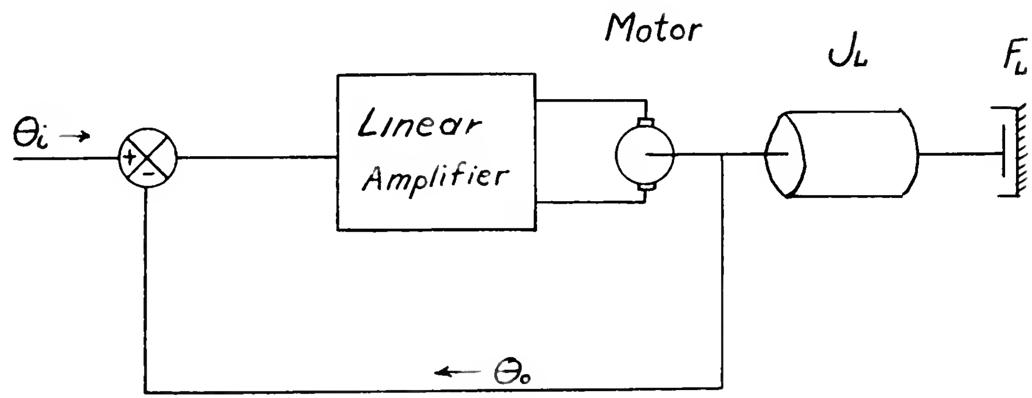


Diagram of a second order, linear position servomechanism

Figure I

for $\frac{K}{J} > \left(\frac{F}{2J}\right)^2$ (5)

$$\theta_o = \Theta\left[1-\epsilon^{-\frac{F}{2J}t}\left(\frac{\frac{F}{2J}}{\sqrt{\frac{K}{J}-\left(\frac{F}{2J}\right)^2}}\sin\sqrt{\frac{K}{J}-\left(\frac{F}{2J}\right)^2}\,t + \cos\sqrt{\frac{K}{J}-\left(\frac{F}{2J}\right)^2}\,t\right)\right]$$

The characteristics of the system are seen to depend on the values of the three parameters J, F and K. If we hold J and F constant and allow K to vary, the equations can be rewritten using new parameters as follows:

let $\frac{F}{2J} = \alpha$

$\frac{K}{J} = \beta^2$

then $\frac{\beta}{\alpha} = \frac{2\sqrt{J}}{F}\sqrt{K}$

for $\beta^2 < \alpha^2$

$$\theta_o = \Theta\left[1-\epsilon^{-\alpha t}\left(\frac{1}{\sqrt{1-\left(\frac{\beta}{\alpha}\right)^2}}\sinh\alpha\sqrt{1-\left(\frac{\beta}{\alpha}\right)^2}\,t + \cosh\alpha\sqrt{1-\left(\frac{\beta}{\alpha}\right)^2}\,t\right)\right] \tag{6}$$

for $\beta^2 = \alpha^2$

$$\theta_o = \Theta\left[1-\epsilon^{-\alpha t}(1+\alpha t)\right] \tag{6a}$$

for $\beta^2 > \alpha^2$

$$\theta_o = \Theta\left[1-\epsilon^{-\alpha t}\left(\frac{1}{\sqrt{\left(\frac{\beta}{\alpha}\right)^2-1}}\sin\alpha\sqrt{\left(\frac{\beta}{\alpha}\right)^2-1}\,t + \cos\alpha\sqrt{\left(\frac{\beta}{\alpha}\right)^2-1}\,t\right)\right] \tag{7}$$

The dimensionless curves in Figure II show the various transient response curves for various values of the parameter

$\frac{\beta}{\alpha}$ and the time constant of the system, where one time constant is:

$$\tau = \frac{1}{\alpha} = \frac{2J}{F} \tag{8}$$

Two terms will be introduced here to aid in explaining the behavior of the system. Rise time will be defined as the time required for the system to first come to the final desired value, regardless of what happens after arriving at this value. Settling time will be defined as the time for the system to come to and stay within some arbitrarily prescribed region of the final value. For this discussion, this region will be defined as the region where the actual value of the system position differs from the final value by an amount of 2% or less.

Examination of the curves of Figure II will show that for values of $\frac{\beta}{\alpha}$ less than unity, the rise time will be infinite and the settling time will be fairly long, that is, greater than six time constants. As the gain is increased, $\frac{\beta}{\alpha}$ increases and at $\frac{\beta}{\alpha}$ equal to one, the system is said to be critically damped. At this value of gain, there is no overshoot, the rise time is still infinite and the settling time is about six time constants.

As the gain is further increased, the rise time and the settling time both decrease, and the system starts to have an overshoot before settling to the final value. As the gain

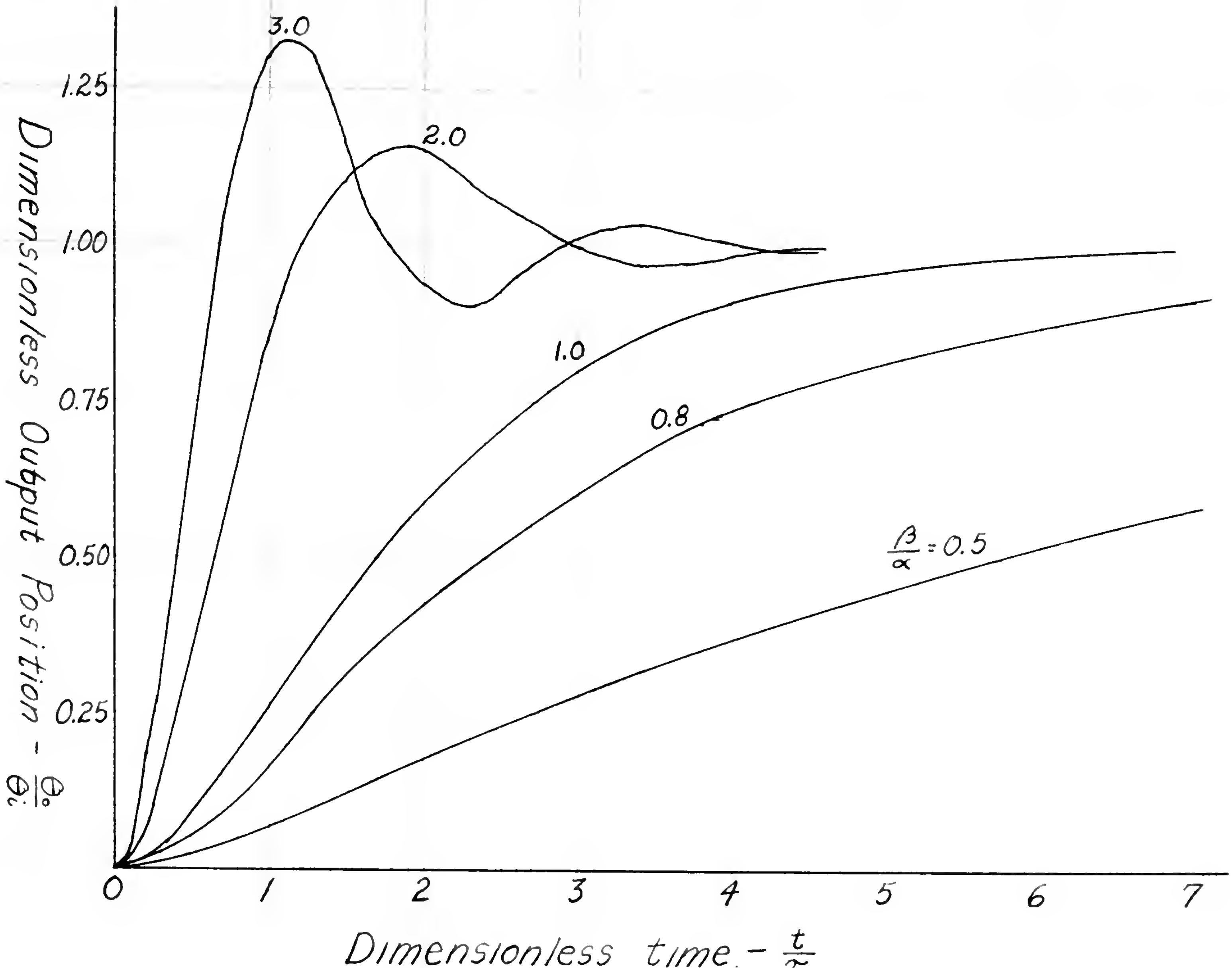

3.0
2.0
1.0
0.8
$\frac{\beta}{\alpha} = 0.5$
1.25
1.00
0.75
0.50
0.25
0
1
2
3
4
5
6
7
Dimensionless Output Position - $\frac{\theta_o}{\theta_i}$
Dimensionless time - $\frac{t}{\tau}$

Figure II

is further increased, the amount of overshoot increases but the rise time decreases. In the limit as the gain approaches infinity, if this were physically possible, the overshoot approaches one hundred percent, the rise time approaches zero but the settling time can never be less than about four time constants.

The fact that the settling time is not less than four time constants can be seen from the equation for the envelope of the damped oscillations, which is:

$$\text{Envelope of } \theta_o = \Theta\left[1 \pm \epsilon^{-\alpha t}\sqrt{\frac{\left(\frac{\beta}{\alpha}\right)^2}{\left(\frac{\beta}{\alpha}\right)^2 - 1}}\right] \tag{9}$$

In the limit as $\frac{\beta}{\alpha}$ approaches infinity, this equation approaches:

$$\Theta\left[1 \pm \epsilon^{-\alpha t}\right] \tag{10}$$

which has a value of $(1 \pm 0.02)\,\Theta$ at $\alpha t = 4$

In practice, the large amounts of overshoot present at high values of gain is undesirable, but a shorter settling time would be very desirable. The ultimate characteristic desired in a servomechanism system would be that it have no overshoot and a zero settling time.

EFFECT OF GAIN CHANGE DURING THE TRANSIENT PERIOD

In the general discussion of the effect of gain on the transient response of a linear system, it was noted that each value of gain had certain advantages and disadvantages associated with it. The low gain systems have the advantage of no overshoot but they have the disadvantages of long rise and settling time. The high gain system has the advantage of shorter rise and settling times but has the disadvantage of having considerable overshoot.

If, during the transient period, some means of varying the gain of the system could be achieved, it is conceivable that the desirable characteristics of the low gain and the high gain systems could be combined to give a system with low overshoot, yet also having a short rise and settling time.

The transient response of such an improved system might be similar to one of the dotted lines shown in Figure III. The solid curves are the same as those in Figure II and are shown on the same set of coordinates for comparison. The transient behavior of these non-linear response curves are such that they combine the rapid acceleration toward the new position of a high gain system, then damp out rapidly when they approach the final value, which is characteristic of the low gain system.

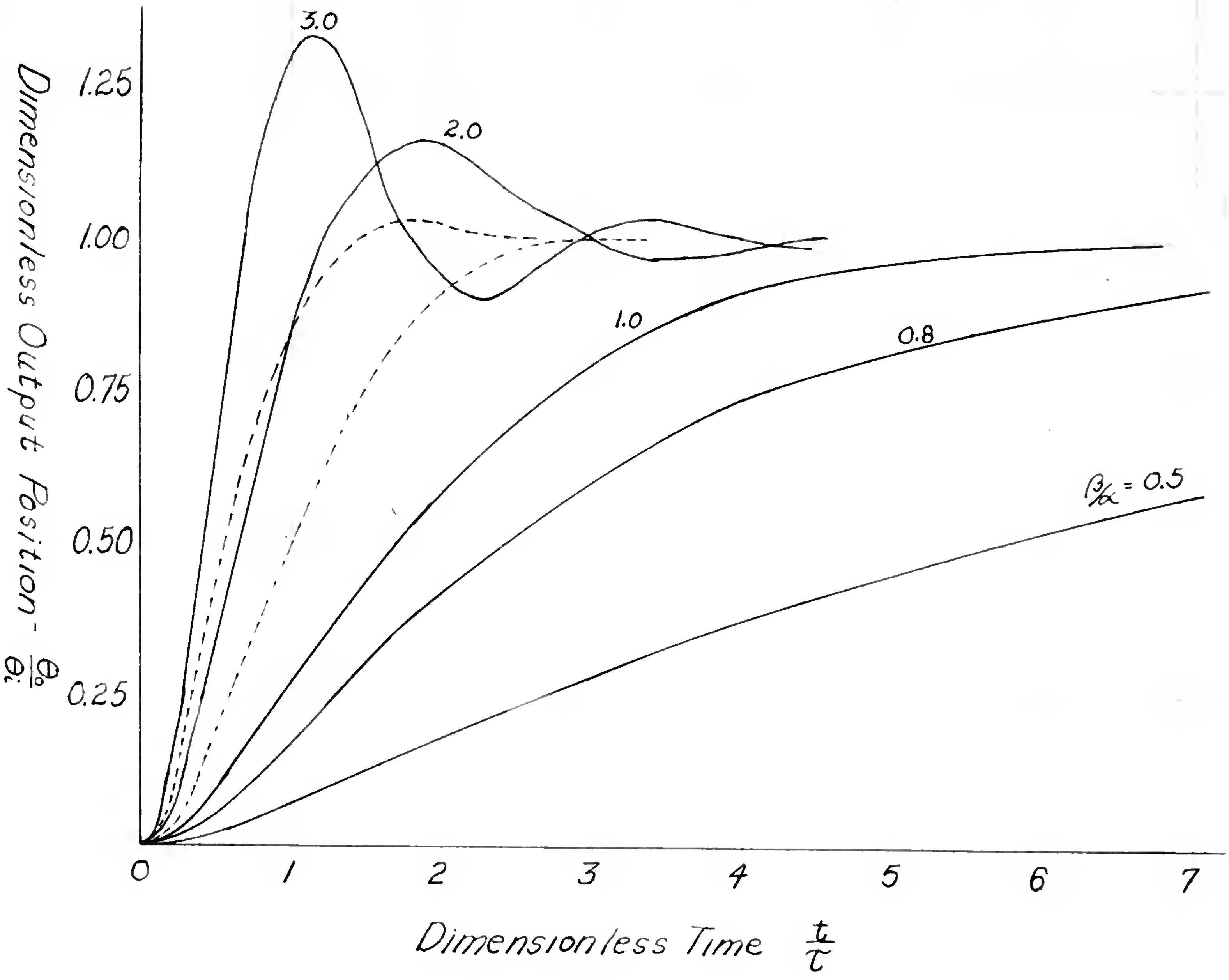

3.0
2.0
1.0
0.8
β/α = 0.5
1.25
1.00
0.75
0.50
0.25
0
1
2
3
4
5
6
7
Dimensionless Output Position - θo/θi
Dimensionless Time t/τ

Figure III

CHOICE OF NON-LINEAR GAIN CHARACTERISTICS

Any amplifier that has an output which is not directly proportional to the input is considered to be a non-linear amplifier. One of the most common types of non-linear gain in an amplifier is that caused by saturation of the amplifier. In this case, the gain is essentially constant over a certain range of input values, but drops off appreciably as the input increases beyond a certain limit. If this type of amplifier were used, the maximum velocity of the system would be limited to some value determined by the voltage input to the amplifier which caused the amplifier to be overloaded.

This characteristic is immediately apparent as being not desirable because we desire the system to attain a very high velocity while it is in transient from the original position to its new position, in order to reduce the rise time of the system. This desire to obtain rapid acceleration to a high velocity as soon as the error is applied indicates that the system should have a high gain at the instant that the step input of displacement is applied to the system.

To satisfy the condition that the system have very little overshoot would indicate that the system should have low gain, particularly when the error signal has been reduced to near zero by the system coming into correspondence with the command signal.

The above considerations lead to the conclusion that the gain of the system must be high when the error signal is first applied and that the gain must reduce as the system comes into correspondence. Thus, the gain of the non-linear amplifier must be large for large signals and small for small signals.

The limits for the gain variation will have values which depend on the type of system performance desired. It is known that minimum overshoot is one of the prime objectives of the system. This fact would indicate that the system gain should be in the vicinity of critical damping, because higher gains result in damped oscillations, with their accompanying overshoot.

If the minimum gain of the amplifier is set at the value giving the system critical damping, the system will have overshoot before coming to rest because of the high velocity acquired by the system when the gain of the amplifier is high to give the acceleration desired when the error is first applied. It has now been decided that the system should have a non-linear gain which has a value at or slightly below the value needed for critical damping for small error signals and which increases to some higher value, limited, perhaps, only by the maximum amount of acceleration that the physical system can withstand, for large error signals.

The manner in which the gain changes between these limits has not been decided. There are any number of possible shapes

for the gain curve which has a certain minimum value and increases as the signal increases. The only feature known about this curve is that is should rise continuously from a value which is small at small error signals to some larger value.

The basic physical concept is that of changing the gain so that the net transient response is the addition of sections of the transient responses for various fixed gains. For example, Figure IV shows the transient response for a linear system for three values of gain, A, B, and C. The non-linear system would operate at gain A for large errors and thus the first section of the response curve would lie on the linear response curve for gain A as shown.

When the error decreases to some value $\mathcal{E}_2$, the gain is changed from A to B and the next s ction of the response curve for the non-linear gain corresponds approximately to that of the corresponding section for the linear system of a gain B. A similar change is made when the error decreases to $\mathcal{E}_3$, shifting to linear gain C. The result is the non-linear response curve D. It should be noted that this physical analysis is not completely correct, since the velocity of the system at the instant of gain change is not the same as would have existed if the lower gain had been used for the entire preceding time interval.

The three basic shapes of gain versus signal amplitude are the straight line, a curve concave upward or a curve

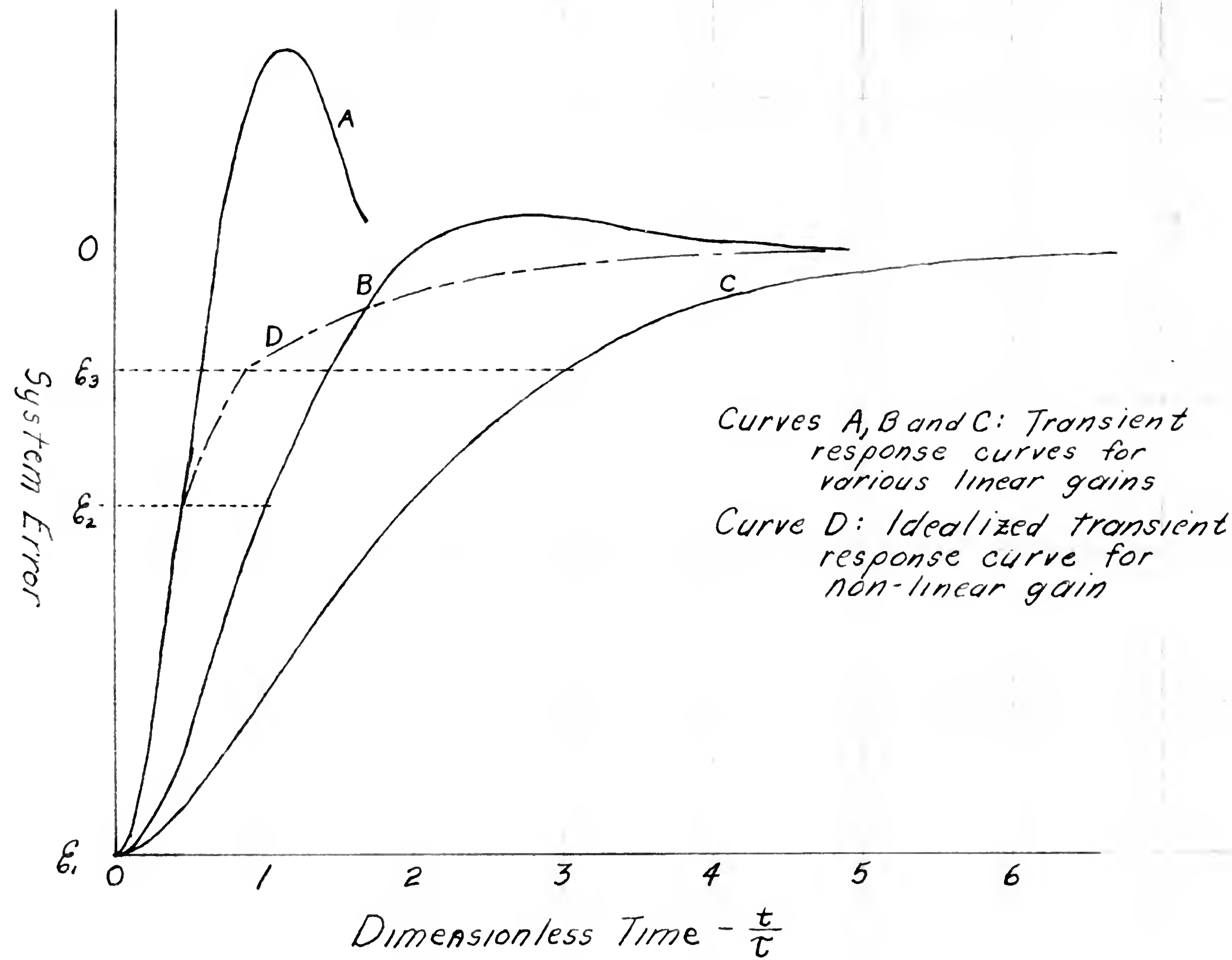

A
B
C
D
0
$\mathcal{E}_3$
$\mathcal{E}_2$
$\mathcal{E}_1$
System Error
0
1
2
3
4
5
6
Dimensionless Time - $\frac{t}{\tau}$
Curves A, B and C: Transient response curves for various linear gains
Curve D: Idealized transient response curve for non-linear gain

Figure IV

concave downward. An attempt to ascertain the effect of the shape of the gain curve will be made when the system is investigated experimentally and transient response curves are obtained.

The low value of gain required for small error signals has disadvantages not heretofore mentioned. These disadvantages are the large velocity lag error and the steady state error in the presence of load torques. The velocity lag error is the displacement error between the position of the command signal and the system output when the command signal is one which has a constant velocity. When the system has reached steady state, the output velocity will be the same as the input velocity, but there will be a difference in their instantaneous positions because for each value of steady state velocity, there must be a control voltage applied to the drive motor, and this control voltage will exist only if there is an error signal caused by a difference in the positions of the input and the output signals. In a system with high gain, the velocity lag error is relatively small, because the control voltage applied to the motor is large for small error, but as the gain is decreased, the error must increase to give the same control voltage to the motor.

If a load torque is applied to the system, this torque will tend to rotate the system output. The system will come to rest only when the motor exerts a torque which is equal in magnitude and opposite in direction to the torque applied by

the load. The steady state (at rest) torque of the motor is directly proportional to the voltage applied to the control winding, and consequently proportional to the error between the command signal and the system output. Here again, the error will be larger in systems having low gain than it would be in a system having high gain.

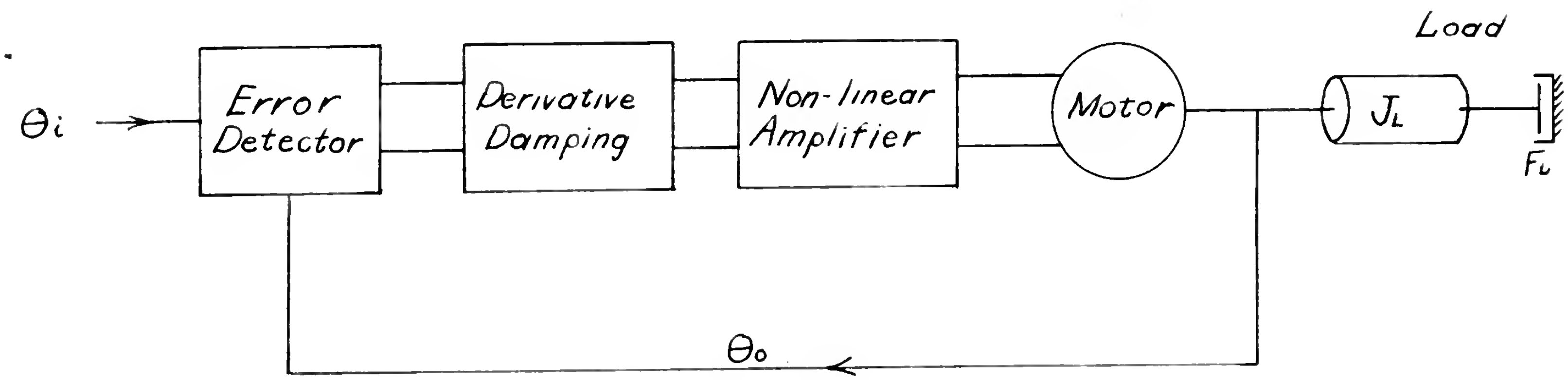
To allow higher gain in the system for the same amount of relative damping, the viscous friction of the system can be increased, thus maintaining the same value for the ratio of $\frac{\beta}{\alpha}$. The same effect as adding viscous friction can also be achieved by the use of derivative damping. This increase in damping has two good effects. One is that it allows higher system gain for the same $\frac{\beta}{\alpha}$ ratio, thus reducing the velocity lag error and load torque error. The other is that the value of α and β each increase, reducing the system time constant. This reduction in the system time constant means that the system will come into correspondence faster, even though it requires the same number of time constants as it did for the smaller values of α and β .

PROPOSED SYSTEM

The function of a position servomechanism is to control some device such that the position of this device will be the same as that dictated at the command station. The device to be controlled can be almost anything conceivable but will, in general, impose some load on the controlling servomechanism. In this theses, the load will be assumed to consist of inertia and viscous friction, designated respectively as J_L and F_L.

The block diagram of the proposed system for investigation is shown in Figure V. The position that is dictated by the command station is shown on the block diagram as θ_i. The actual position of the load is designated as θ_o. The other functions to be performed to make the system complete are those of the error detector, derivative damper, non-linear amplifier and the motor.

The function of the error detector is to measure the position of the load and compare this position with that dictated by the command station, and to provide a signal which is a function of the difference between the actual position and the command position. The signal is then fed to the derivative damper, which changes the signal from a function of the positions only, to one which is also a function of the derivative of the positions, or to the velocities. The result of this damper is to increase the coefficient of the velocity term in the system equation, thus increasing the damping constant, α , for the system.



Block Diagram of Proposed System

Figure V

An amplifier is added to the system to change the power level from the error detector, or derivative damper, if used, to a higher power level. The power output from the error detector is usually quite small, on the order of a few watts or less, while the power required to position the load may be several hundred horsepower. Such large power is not contemplated in this investigation, but a power level increase will be needed. The amplifier commonly used is linear, that is, it produces a power output which is considerably larger than the power input, but is directly proportional to it. Since the center of this investigation is the effect of non-linear gain, the amplifier will be non-linear.

The motor is necessary to convert the electrical power output of the amplifier to mechanical for positioning the load.

CHOICE OF COMPONENTS

The functions of the various components of a servo system are fairly well defined but there are numerous ways and means to accomplish these functions. The most apparent difference in components is the fact that some are applicable only to systems which use an a.c. voltage as an indication of the system error, others are applicable only to systems which use a d.c. voltage as the indication of error, and still others can be used in either type of system. The selection of which actual components will be chosen to perform the various functions depends upon many factors, among others, power supply voltages available, power handling requirements, physical size, inertia and friction effects, and relative ability to perform the function necessary. The following discussion deals with the reasons for selecting or not selecting certain components used in the system which was built for investigation.

Since the effect of non-linear gain is the center of the investigation, the non-linear amplifier was the first component to be selected, then the rest of the system was built up around it. When it was decided that the amplifier should have the characteristic that the gain increases with increasing error signal, the circuit commonly known as a volume expander was first considered. This circuit rectifies a portion of the a.c. signal and uses this rectified

voltage as a variable bias to the grid of a variable mu vacuum tube. Examination of this circuit showed that it had several undesirable properties. First, the application of the bias voltage to the grid of the amplifier stage introduces a transient into the signal circuit known as thump. This thump can be minimized by proper circuit design but is an objectionable feature. The second property is that the action of the variable gain is not instantaneous. This means that the gain is not only a function of the amplitude of the signal but also of the rate of change on the signal amplitude. This fact in itself may not be objectionable, but the fact that the relationship between gain and rate of change of signal amplitude is not the same for increasing signals as it is for decreasing signals is objectionable. The third, and probably most important feature is that there is very little control of the shape of the gain vs signal amplitude characteristics of the amplifier.

The next method considered for producing a circuit that would increase gain with increased signal amplitude was the use of diodes in a limiter circuit. This circuit uses diodes to shunt out certain resistors in a voltage divider network, thereby changing the portion of the input that is available at the output terminals. This circuit is applicable primarily to the d.c. error voltage type of electrical system. This limiter circuit can be adjusted over a very wide range of operating characteristics and has no time lag as has the

volume expander circuit mentioned above. The limiter was chosen for use in this thesis because of these characteristics. A diagram of the circuit is shown in Figure IX. A more detailed description of this circuit will be found in Appendix A.

This circuit could be used in an a.c. system, but it would introduce considerable third and higher odd order harmonics into the carrier voltage wave. The a.c. motor would respond primarily to the amplitude of the fundamental component of the carrier and the value of this would vary with the amount of non-linearity and cannot be readily determined for analysis. For this reason, it was decided that the error channel should be d.c.

The next component considered was the error detector. Since the non-linear circuit was to operate with a d.c. signal, a potentiometer was chosen to provide this signal. The voltage applied to the potentiometers was twelve volts, to prevent overloading and burning out the potentiometers. This provided a maximum error signal of about 6 volts, but the actual amplitude based on the amount of travel of the slider was on the order of one volt.

The non-linear circuit was designed for a signal level of about ten volts input, so a d.c. amplifier was built and inserted between the error detector and the non-linear circuit. This circuit is shown in Figure VIII, and is further described in Appendix A.

A drive motor was needed, and since the system has been designed to use a d.c. error signal up to this point, it was decided to carry this throughout the system. A separately excited, one quarter horsepower, 250 volt d.c. motor was installed.

Since considerable power was required to drive this motor, a General Electric Model 5AM45DB11A, 250 volt, one ampere amplidyne generator was used as the source of power. The split fields of the amplidyne were excited from two 6L6 vacuum tubes in push pull, the current in the fields bucking each other so that there was no net excitation when the currents in the tubes were the same.

This completed the original selection of components and preliminary tests were made on the complete system. These tests showed that this system would not be satisfactory for obtaining results as to the effect of the non-linear gain. The numerous time constants encountered in the amplidyne and motor circuits increased the order of the system from a second order system, described earlier, to about a fourth or fifth order system. Systems of this type are generally unstable at even moderate gains, unless special precautions are taken to adequately compensate for the large phase shift through the system. When the gain was reduced sufficiently to prevent instability in the system, the static friction of the motor, being primarily in the brushes, was of such large magnitude compared to the viscous friction that the system would not

follow the command signal. This inability to follow the command signal rendered the system useless for test purposes.

The next step in the construction of the system was to find a motor that had very little static friction. No suitable d.c. motor could be found, so the possibility of using a two phase a.c. motor was investigated. A two phase motor, particularly one in which the rotor is mounted on ball bearings, has very little static friction due to the absence of brushes or other sliding contacts. A Diehl Model FPE-49-13-1, ten watt, 115 volt, 400 cycle two phase motor was chosen for the job because of its relatively linear operating characteristics. The main field of this motor was supplied directly from the 115 volt, 400 cycle supply available in the laboratory. Since the error signal was a d.c. signal, and the non-linear circuit was designed to be used on d.c., it was necessary to provide a means of changing the d.c. error signal to a 400 cycle error signal for application to the control field of the servo motor. The requirements that must be met by such a device are that the amplitude of the 400 cycle signal must be in proportion to the magnitude of the d.c. signal and there must be a 180 degree phase shift of the 400 cycle signal when the d.c. signal changes from plus to minus. This last requirement is necessary in order that the motor will reverse when the error reverses.

There are at least two possible methods for accomplishing this. The first method of doing this is to apply the d.c.

error signal to the field of a 400 cycle alternator, running at synchronous speed and properly phased with the voltage applied to the main field of the motor. The voltage applied by the alternator to the control field is properly phased if it leads the main field voltage by 90 degrees for a given polarity of the error signal and if it lags by 90 degrees for the opposite polarity of the error signal. This method was considered impractical because of the cost of construction and the relatively large physical size of the converter. The inductance in the field of the alternator would also add an extra time constant to the system, which is considered undesirable.

A second method of accomplishing the same result is to use a device known as a chopper. This device was the one actually used and the circuit is shown in Figure X. The principle of operation is that the vibrator contact vibrates at a frequency which is the same as the frequency applied to the vibrator coil. This vibrating contact connects first one end of the primary of the transformer to the input terminal a, then the other end of the transformer is connected to the same terminal, the center tap of the primary being connected to the other input terminal b. This action reverses the flux through the core of the transformer and results in a voltage at the output terminal which has the same frequency as that of the vibrating contactor. The output is essentially a square wave if a high quality transformer is used. In the

circuit constructed, a Stevens-Arnold type 268 chopper and a United Transformer Corporation type A-26 transformer were used. It is realized that this circuit also introduces harmonics into the 400 cycle carrier, but the fundamental content is proportional to the d.c. input, and is not affected by the amount of non-linearity in the circuit. This 400 cycle voltage was then fed through an electronic power amplifier to obtain sufficient power to the control field of the motor.

A means of controlling the phase relationship of the voltage applied to the control field, with respect to the voltage applied to the main field of the motor, was required. A resistance-capacitance network can be used in the signal circuit but, unless great care is taken in choosing components with high accuracy, the network will have different insertion losses for different phase shift settings. Another method to accomplish this phase shift control is to control the phase of the voltage applied to the vibrator coil in the chopper.

To accomplish this phase shift control in the vibrator circuit, three phase, 115 volt, 400 cycle voltage, obtained from a Scott connected transformer which received excitation from the two phase supply in the laboratory, was applied to the three phase winding on the stator of a synchro generator, as shown in Figure XIII. The rotor output has a magnitude which is independent of the angular position of the rotor, but the phase relationship of the rotor voltage can be continuously varied through 360 degrees by rotating the rotor

of the synchro generator. This rotor output was fed through a step-down transformer to supply the necessary six volts to the vibrator coil and the phase relationship varied by rotating the position of the rotor of the synchro generator.

The system then consisted of the following components: a d.c. potentiometer error detector, a d.c. amplifier, the non-linear circuit, a chopper, an a.c. power amplifier, and the two phase motor.

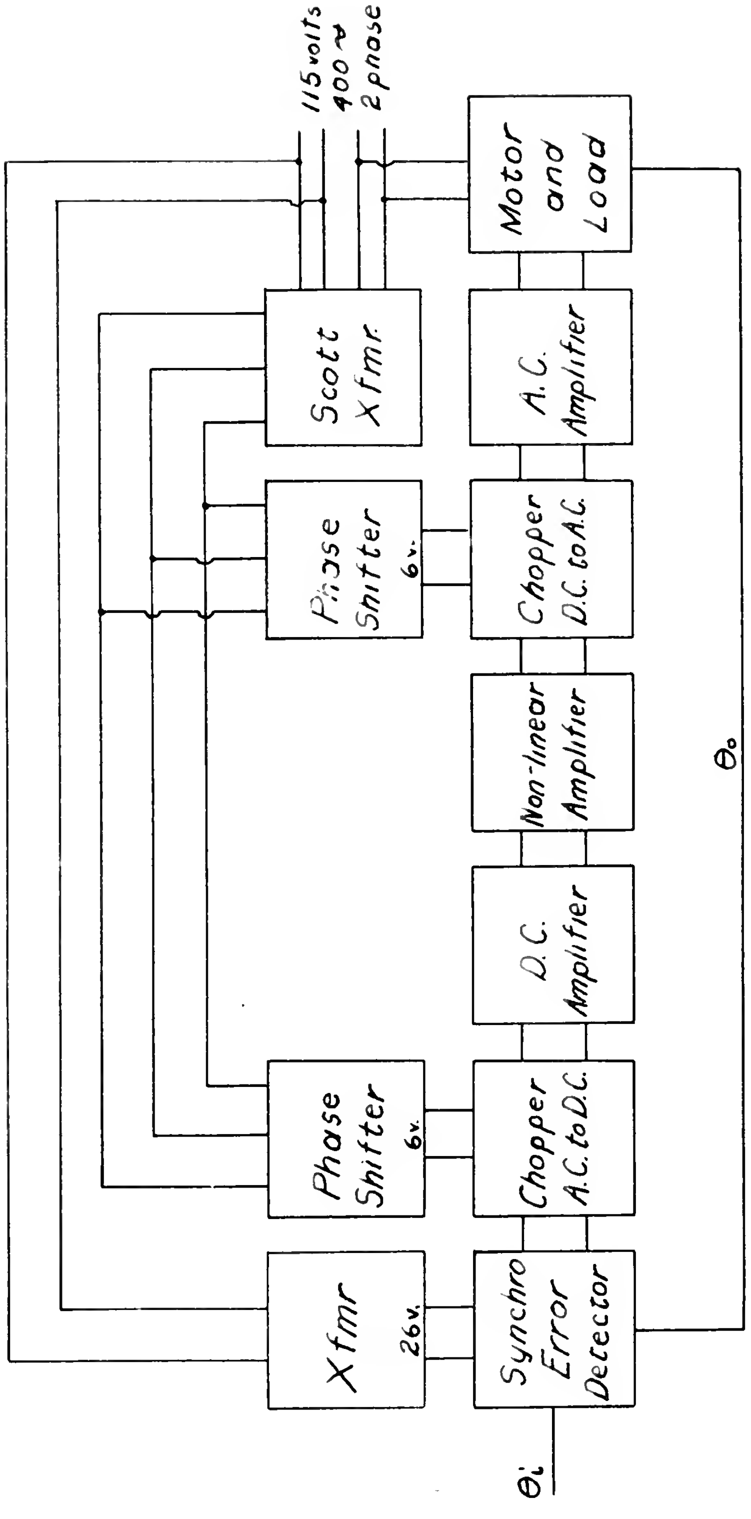
Tests on this system showed improvement over the system using the d.c. motor in regards to static friction, however some static friction was still present. At gains in the region where the system was underdamped, the effect of this static friction was not too apparent, but as the gain was reduced into the overdamped region, the effect became very noticeable. A further reduction in the static friction was indicated as being necessary. Investigation of the system showed that the greatest friction was in the slider of the potentiometer error detector.

The chopper circuit previously discussed for converting a d.c. error signal to an a.c. error signal is also suitable for converting an a.c. signal to a d.c. signal if the input terminals and the output terminals are interchanged, and a small filter added to reduce the ripple in the d.c. signal. This feature of the chopper made the use of an a.c., synchro error detector possible, still maintaining a d.c. error signal through the non-linear circuit. Two Pioneer type AY-201-

3-B, 26 volt, 400 cycle synchros were used for the error detector. The friction in these units was far less than the friction in the potentiometers previously used.

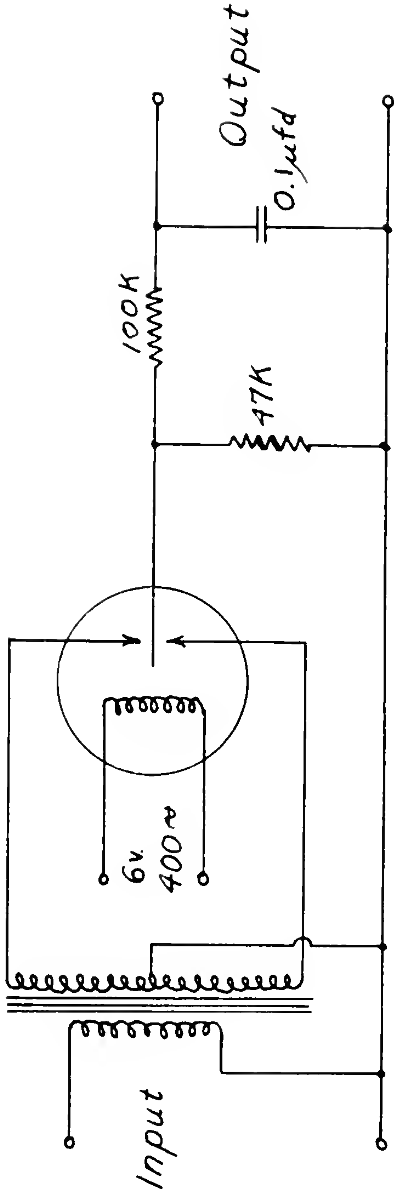
To further minimize the effect of the static friction, additional viscous friction was added, permitting higher gain <u>for the same amount of relative damping</u>. <u>This higher gain</u> reduced the steady state error of the system, but due to power handling limitations of the motor and power amplifier, a sufficient amount of gain and viscous friction to make the effect of static friction negligible could not be attained. The viscous damping was provided by gearing an induction motor to the drive motor, the fields of the induction motor being excited by direct current. An induction motor so excited produces a retarding torque which is very nearly proportional to the velocity of the rotor, and consequently to $\frac{d\theta_o}{dt}$.

The final selection of components resulted in a system which consisted of a synchro error detector, a chopper for converting the a.c. error signal to a d.c. error signal, a linear d.c. amplifier, a chopper to convert the d.c. error signal back to an a.c. error signal, an a.c. amplifier and a two phase motor, whose load was the viscous friction supplied by the induction motor with d.c. in the field windings and the combined inertias of the motors, gears and error detector. The derivative damper mentioned in the proposed system was not used because variable damping was available by varying the amount of d.c. in the winding of the induction motor. The block diagram for this system is shown in Figure VI, with circuit diagrams of the various components shown in Figures VII through XIII.



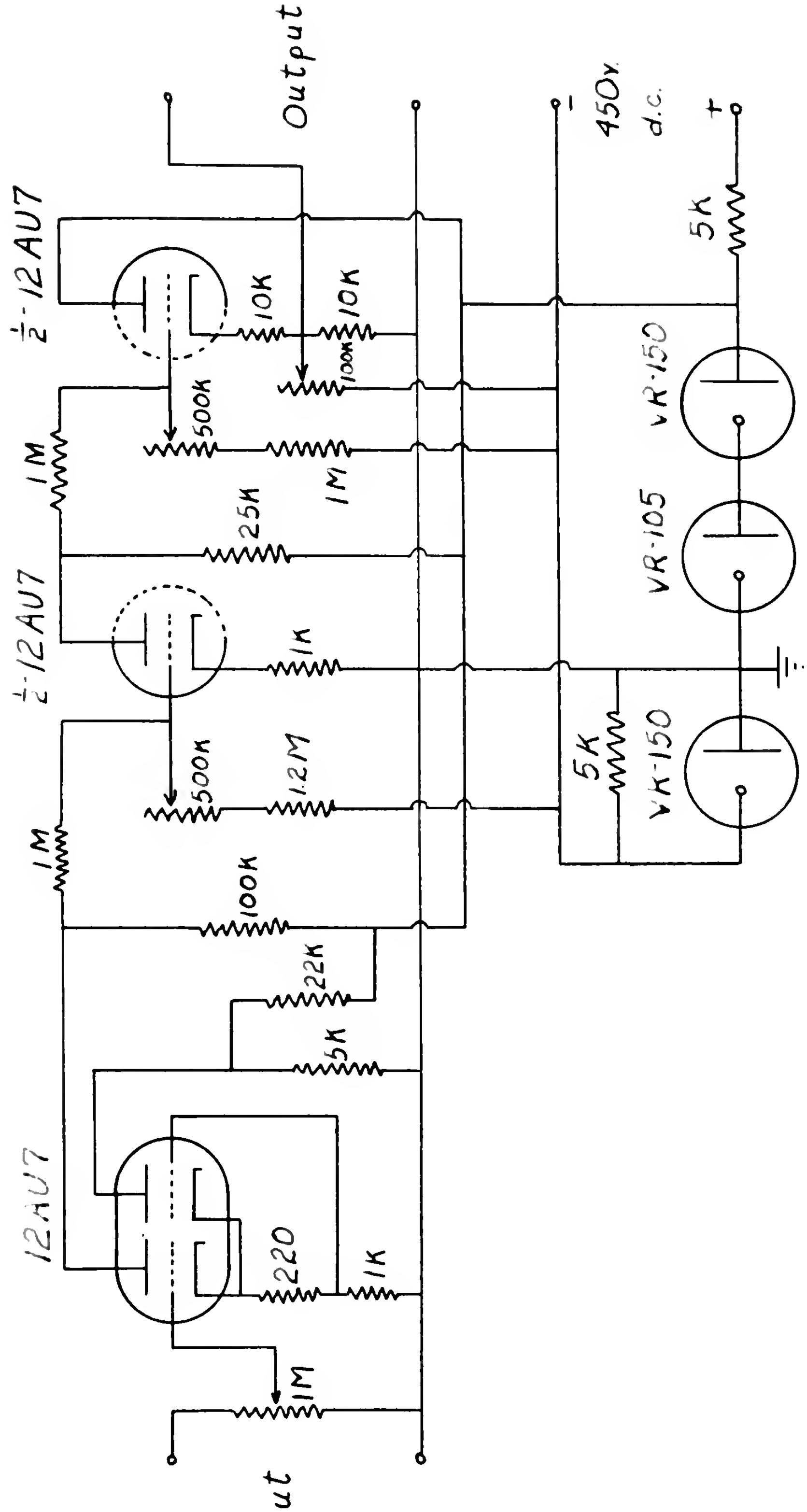
Block Diagram of Final System

Figure VI



Chopper Circuit - a.c to d.c

Figure VII



Schematic of d.c. amplifier

Figure VIII

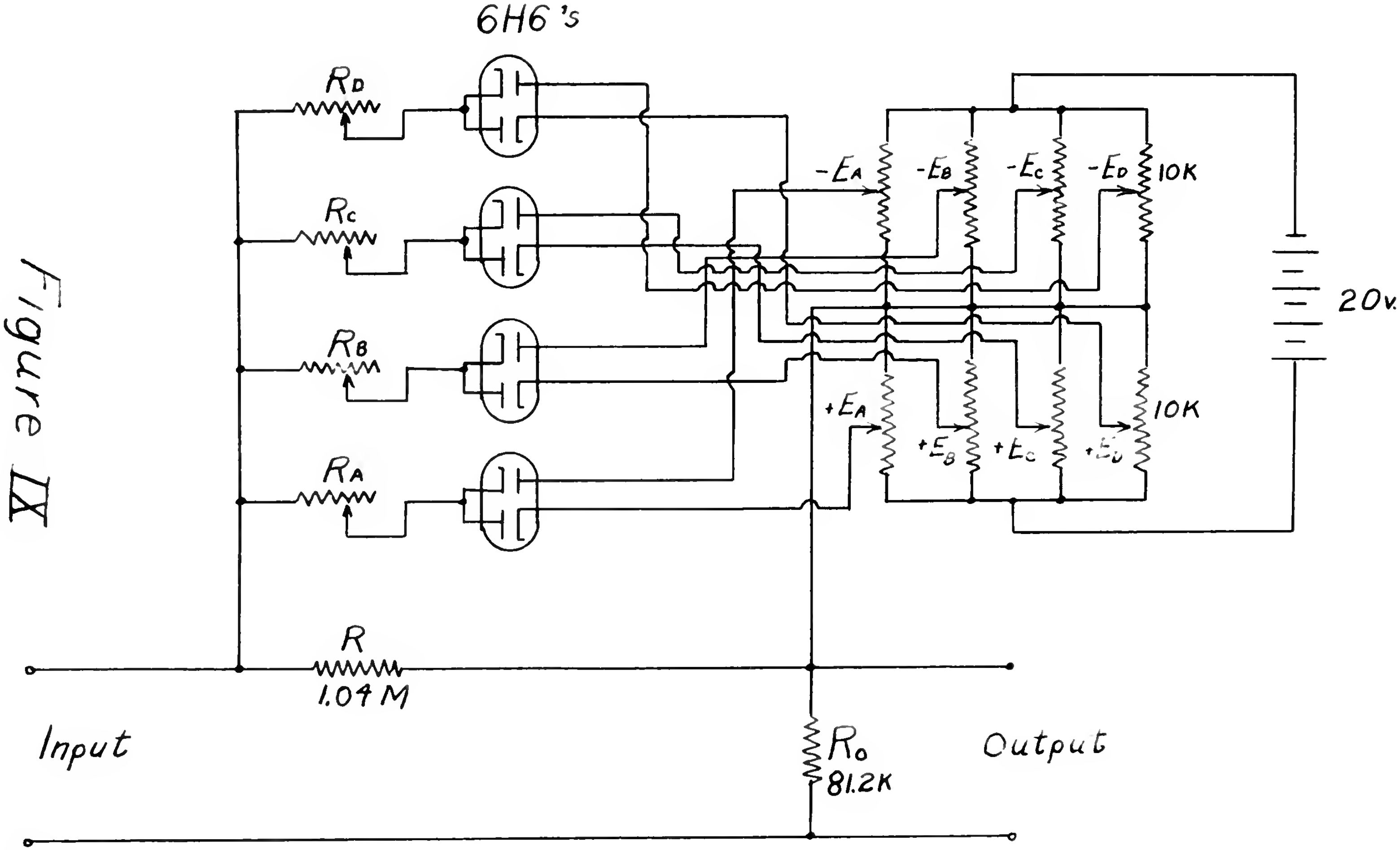
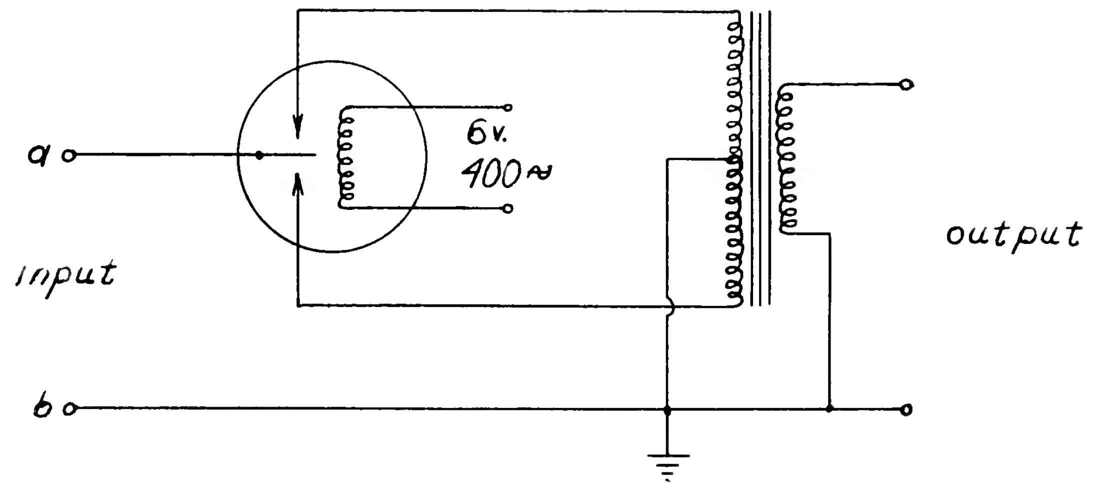


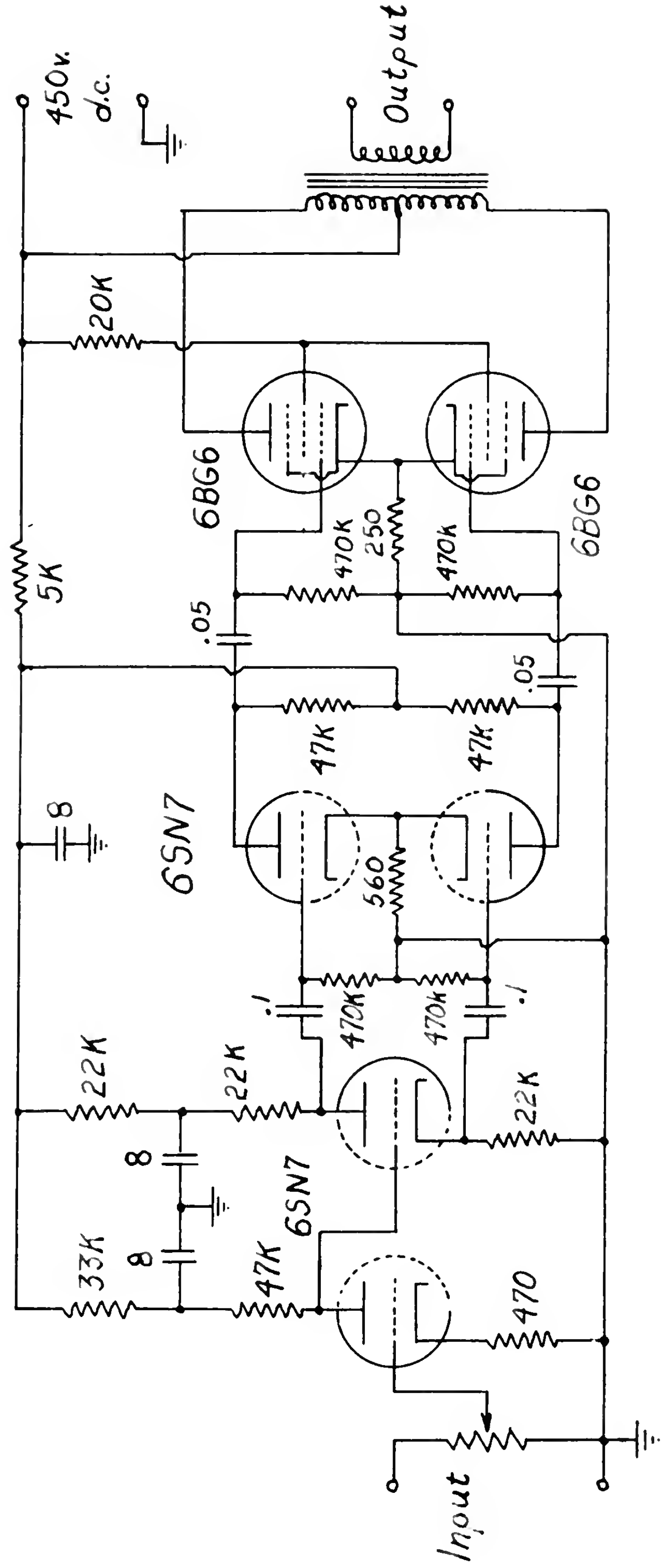
Figure IX

Non-linear Circuit



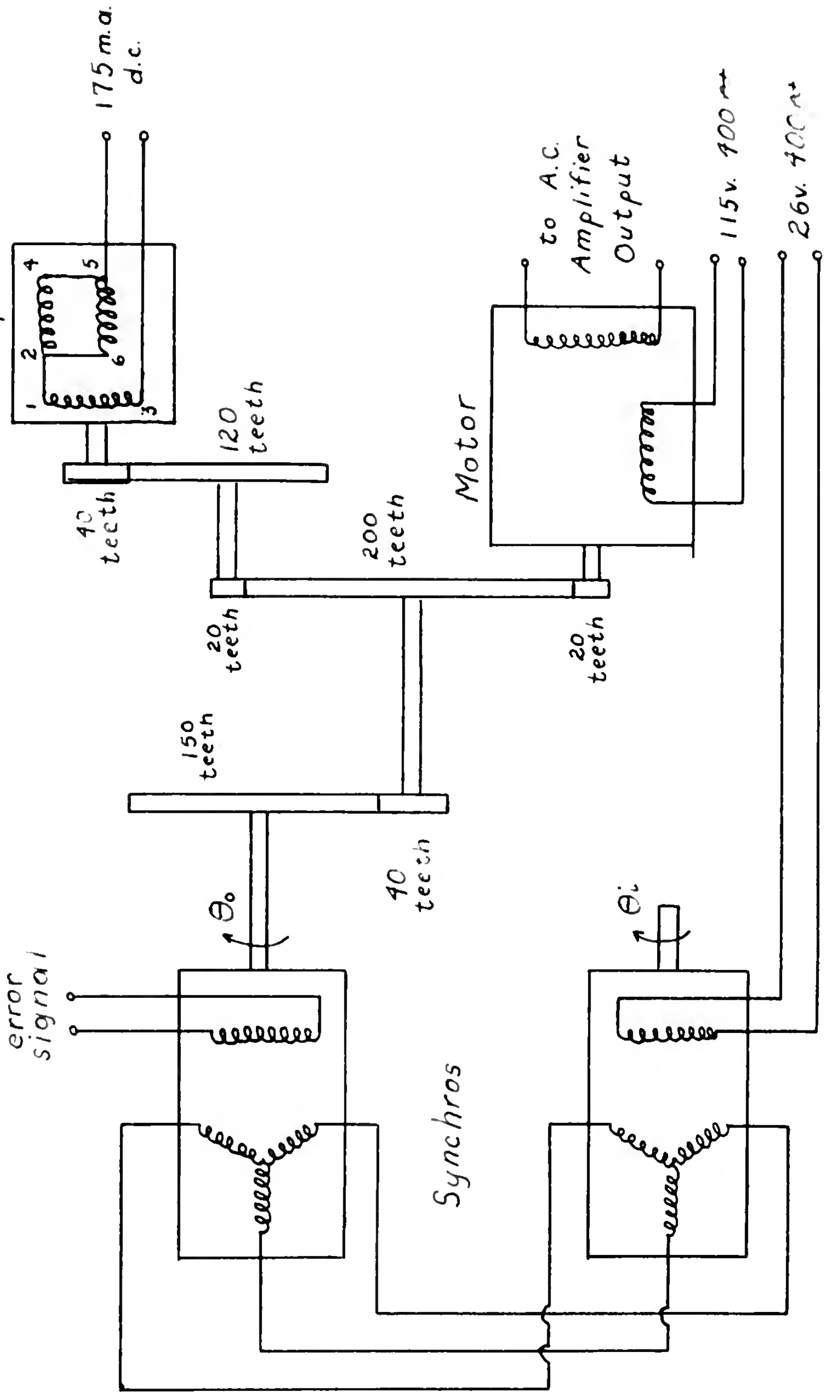
Chopper Circuit - d.c. to a.c.

Figure X



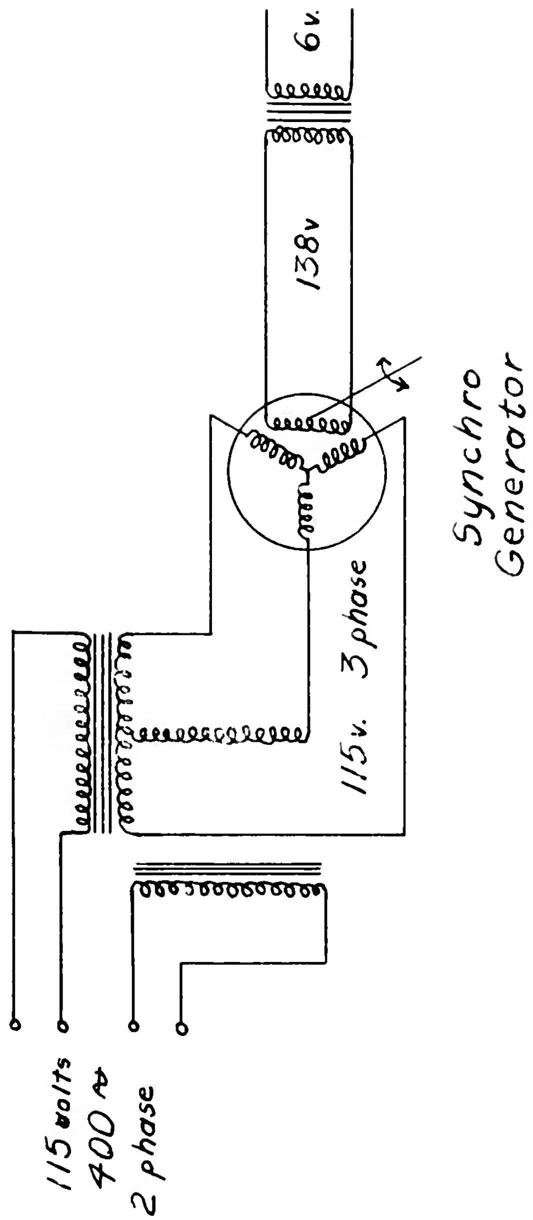
Schematic of A.C. amplifier

Figure XI



Motor, Load and Error Detector

Figure XII



Scott Transformer and Phase Shifter

Figure XIII

SAMPLES OF EXPERIMENTAL DATA FROM THE PHYSICAL SYSTEM

Two sample curves taken on the physical system constructed for use in this thesis are shown in Figures XIV and XV, which were taken with the system operating with linear gain. The purpose of taking curves of the system response using linear gain was to establish a reference of comparison to show the effect of non-linear gain. Figure XIV was taken with the linear gain adjusted to give a system which was slightly overdamped. The value of the command signal is indicated on this curve to show steady state error of the system in the presence of static friction. The initial jump in the transient response curve was probably caused by backlash in the gears.

Figure XV was taken with the system adjusted to give a slightly underdamped system. Again, the effect of backlash in the gearing can be noted. The system has very little steady state error due to the higher gain used, but the distortion or flattening of the overshoot peak was apparently caused by the system sticking when it came to rest, then finally coming back to the desired value.

Transient response curves were made using various types of non-linear gain characteristics and some effect was noted, but the distortion of the linear transient response curves was of such a magnitude that it was very difficult to differentiate between the effects of the non-linear gain and

the distortion caused by the static friction and backlash. It was felt that any results obtained from such data would be inconclusive so this method of investigating the effect of non-linear gain was not pursued any further. The use of the analog computer was felt to be the best approach to the solution.

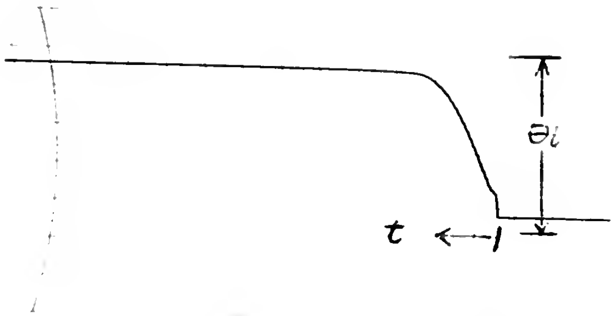


Figure XIV

Transient Response of Physical System
Overdamped

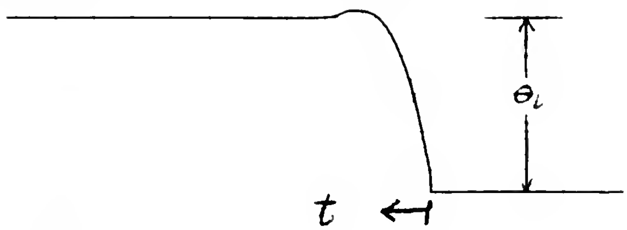


Figure XV

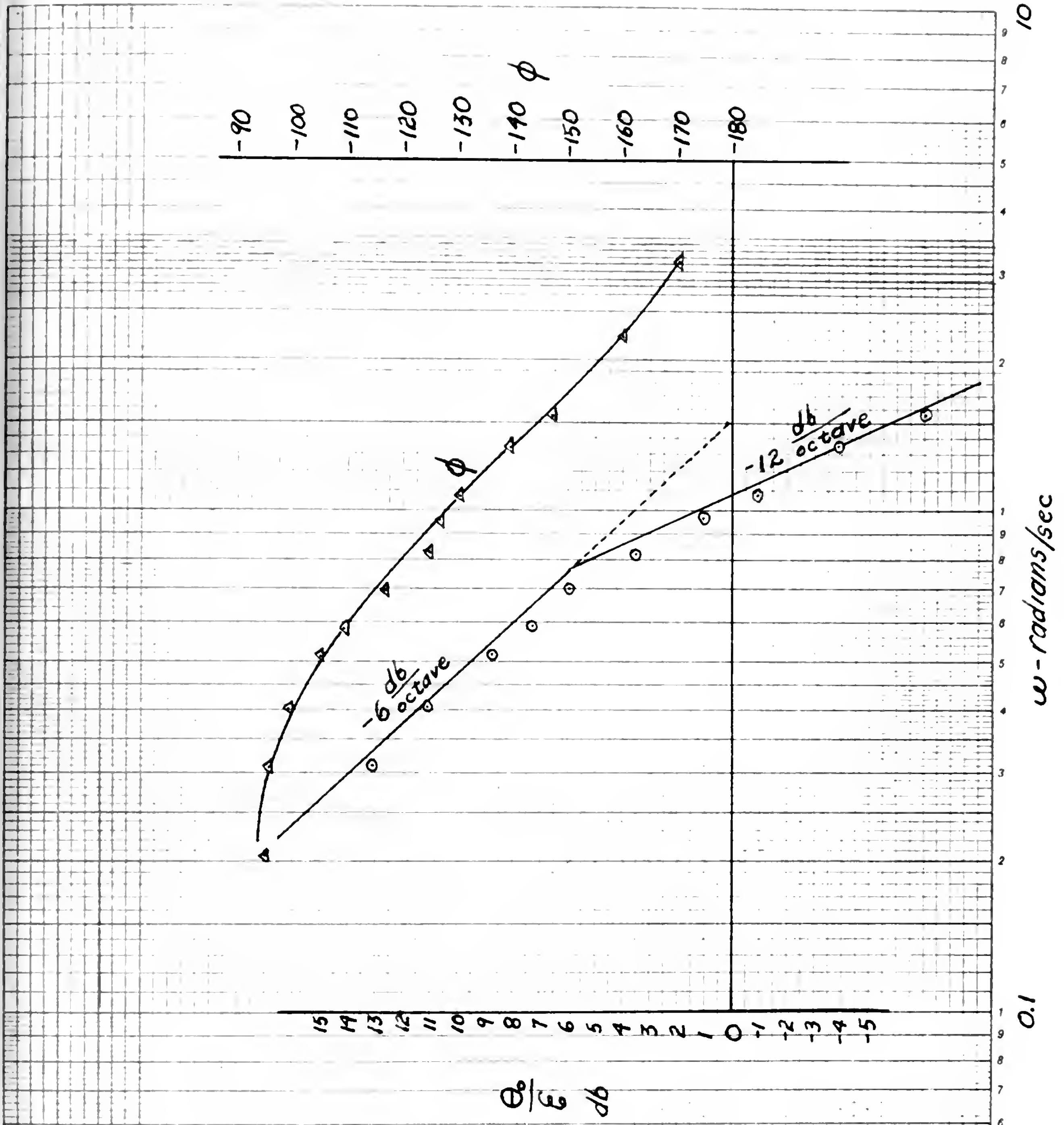
Transient Response of Physical System
Underdamped

USE OF THE ANALOG COMPUTER

When it was found that no accurate conclusions could be made from the transient curves obtained from the actual system set upffor investigation, it was decided that the analog computer could be used to study the effect of non-linear gain on the performance of a servo system, without encountering the difficulties and inaccuracies encountered in the actual system built up.

It was desired that the system studied on the analog computer should be a system similar to the one actually constructed, so a frequency response run was made on the actual system. From the frequency response data, the transfer function of the system was found and the curves for this transfer function are shown in Figure XVI. The fact that the phase shift of the system is between minus ninety and minus 180 degrees indicates that the system is a second order system. The angular frequency at which the asymptote for the response curve changes from minus six decibels per octave to minus twelve decibels per octave is 7.6 radians per second and the intercept of the minus six decibel per octave line with the zero decibel line is at 15.7 radians per second. These values give the transfer function equation:

$$\frac{\theta_o}{\varepsilon} = \frac{15.7}{j\omega\left(\frac{1}{7.6}j\omega+1\right)} = \frac{15.7}{j\omega\left(0.131j\omega+1\right)} \qquad (11)$$



Transfer Function Plot

Figure XVI

Analysis of a second order servo system shows that the transfer function of the system is:

$$\frac{\theta_o}{\varepsilon} = \frac{\frac{K}{F}}{j\omega\left(\frac{J}{F}j\omega+1\right)} \tag{12}$$

Thus the values for $\frac{K}{F}$ and $\frac{J}{F}$ are known, and are:

$$\frac{K}{F} = 15.7 \qquad \frac{J}{F} = 0.131$$

To set a problem up on the analog computer, the differential equation of the system must be known. The differential equation equation of the second order servo system is:

$$J\frac{d^2\theta_o}{dt^2} + F\frac{d\theta_o}{dt} + K\theta_o = K\theta_i \tag{1}$$

Rearranging to a more convenient form for use in the computer we get:

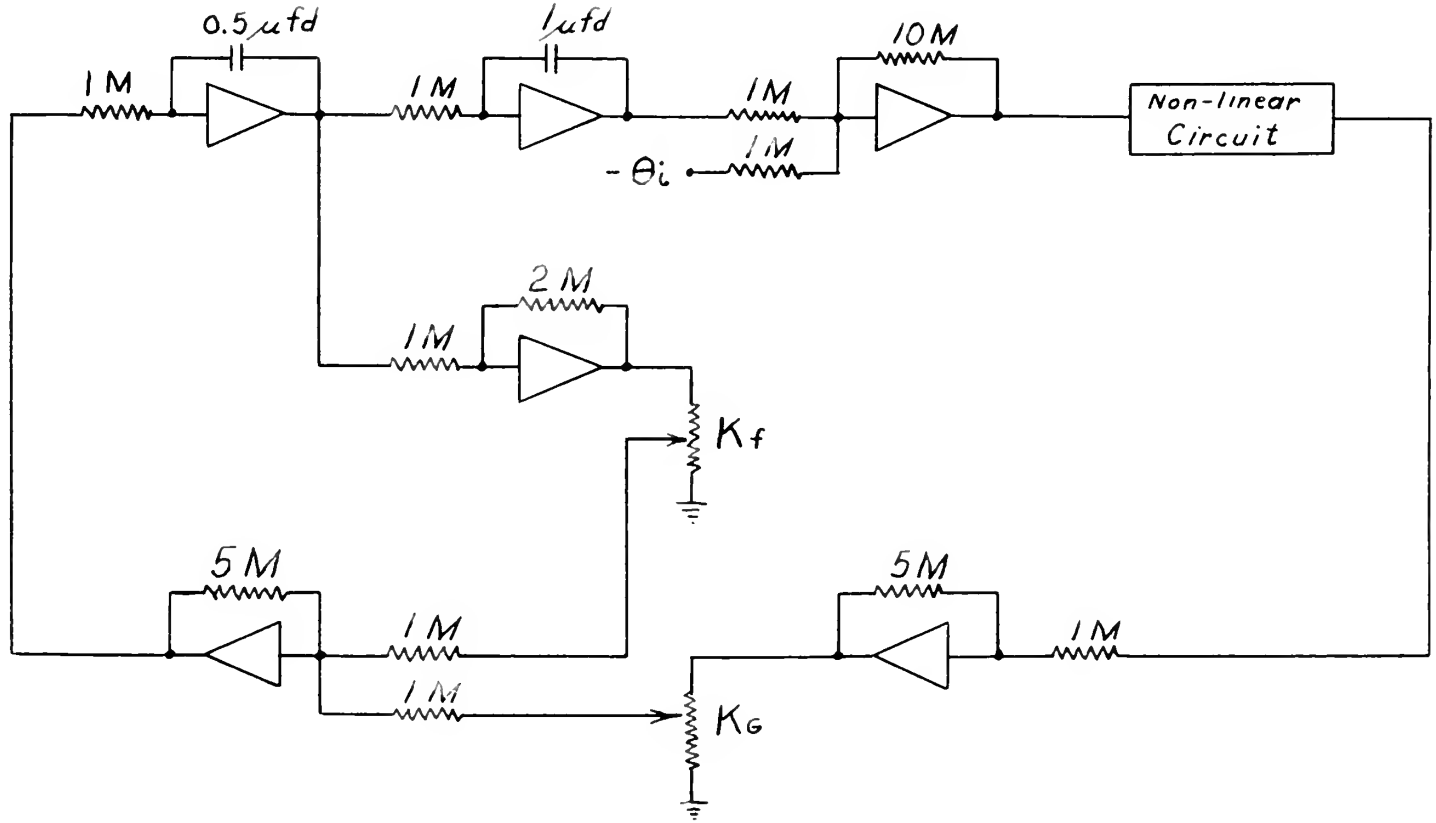
$$\frac{d^2\theta_o}{dt^2} = \frac{K}{J}(\theta_i - \theta_o) - \frac{F}{J}\frac{d\theta_o}{dt} \tag{13}$$

Since F and J are to remain constant in the actual system, while K is to be a variable, the value of $\frac{F}{J}$ can be substituted, giving:

$$\frac{d^2\theta_o}{dt^2} = \frac{K}{J}(\theta_i - \theta_o) - 7.6\frac{d\theta_o}{dt} \tag{14}$$

The computer circuit corresponding to this equation, including the effect of the non-linear amplifier, is shown in Figure XVII. The potentiometer marked K_f was set to give the

value of $\frac{F}{J}$ of 7.6, while the potentiometer marked K_g was adjusted to give the desired value for $\frac{K}{J}$. The insertion of the non-linear circuit, as shown, made the circuit applicable to a system with non-linear gain. With the non-linear circuit in use, the setting on the potentiometer K_g determined the system gain at zero system error, and the non-linear circuit then increased the gain for increasing values of error in a manner which depends on the settings in the non-linear circuit. To find the shape of the non-linear characteristic it is only necessary to calculate the gain of the non-linear circuit and multiply by the linear gain of the remainder of the circuit determined by the setting of potentiometer K_g. The method used to compute the characteristic of the non-linear circuit is to be found in Appendix A.



Computer Circuit
Figure XVII

SAMPLES OF EXPERIMENTAL DATA FROM THE ANALOG COMPUTER

The first information obtained from the analog computer was a series of transient response curves for linear gain to give a set of values as the basis of comparison to evaluate the performance of the system when non-linear gain was used. In all cases throughout this investigation, the value of $\frac{F}{J}$ was set at 7.6 and held constant. This value of $\frac{F}{J}$ gave a system time constant of 0.263 seconds. The values that were of particular interest in establishing a basis of comparison were rise time, settling time, and percent overshoot. Figure XVIII shows a transient response curve for the system when $\frac{K}{J}$ was set at 100. The abscissa of the curve is time and the ordinate is percent of final value, and the scale factors are such that each division of time represents 0.2 seconds and each large division of the ordinate is twenty percent (each sub-division represents four percent) of the final value. The value of abscissa or ordinate, as appropriate, which indicate the value of the rise time, settling time and percent overshoot are indicated on the curve.

A transient curve for the system with linear gain such that the system is overdamped, $\frac{K}{J}$ is equal to 10, is shown in Figure XIX. There is no overshoot on this curve and the rise time is infinite, so the value of settling time is the only value of significance to be obtained from the curve.

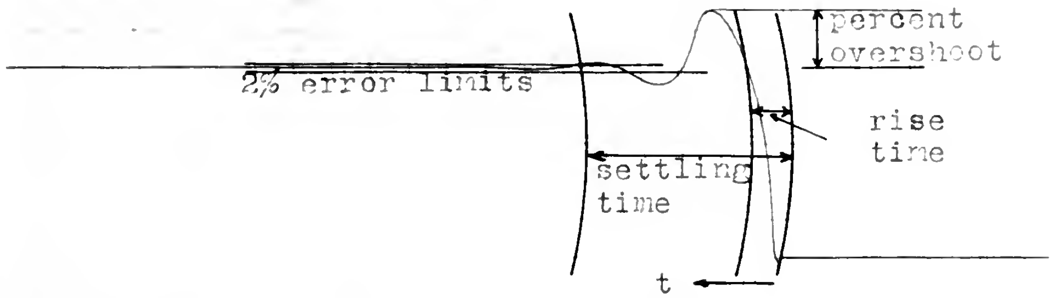


Figure XVIII

Underdamped Transient Response, Linear Gain

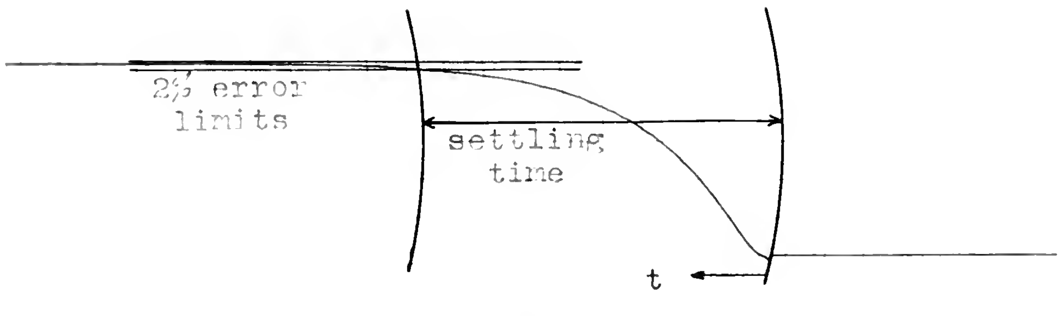
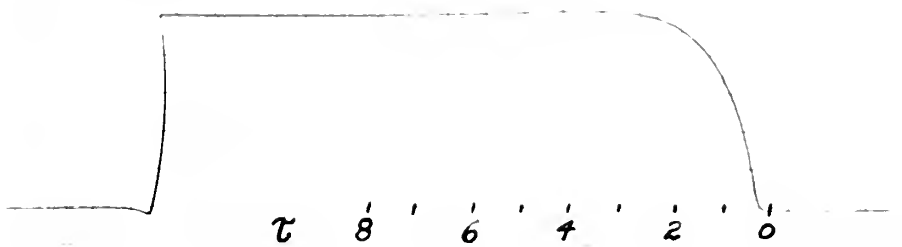


Figure XIX

Overdamped Transient Response, Linear Gain

Tabulated values of rise time, settling time and percent overshoot are given in Figure XXVIII, Appendix B, and curves of these values versus $\frac{K}{J}$ and $\frac{\beta}{\alpha}$ are shown in Figures XXIX, XXXI and XXX respectively.

A typical gain vs system error and the resulting transient response curve is shown in Figure XX. Additional curves of gain vs error characteristics and associated transient response curves are shown Figures XXXII through XLII, Appendix B. The value of Θ_i used when obtaining the transient response is indicated on the appropriate figure.



θ_i = 1.0 volt

Overshoot - 3 %

Rise Time - 0.6 sec - 2.3τ

Settling Time - 1.4 sec - 5.3τ

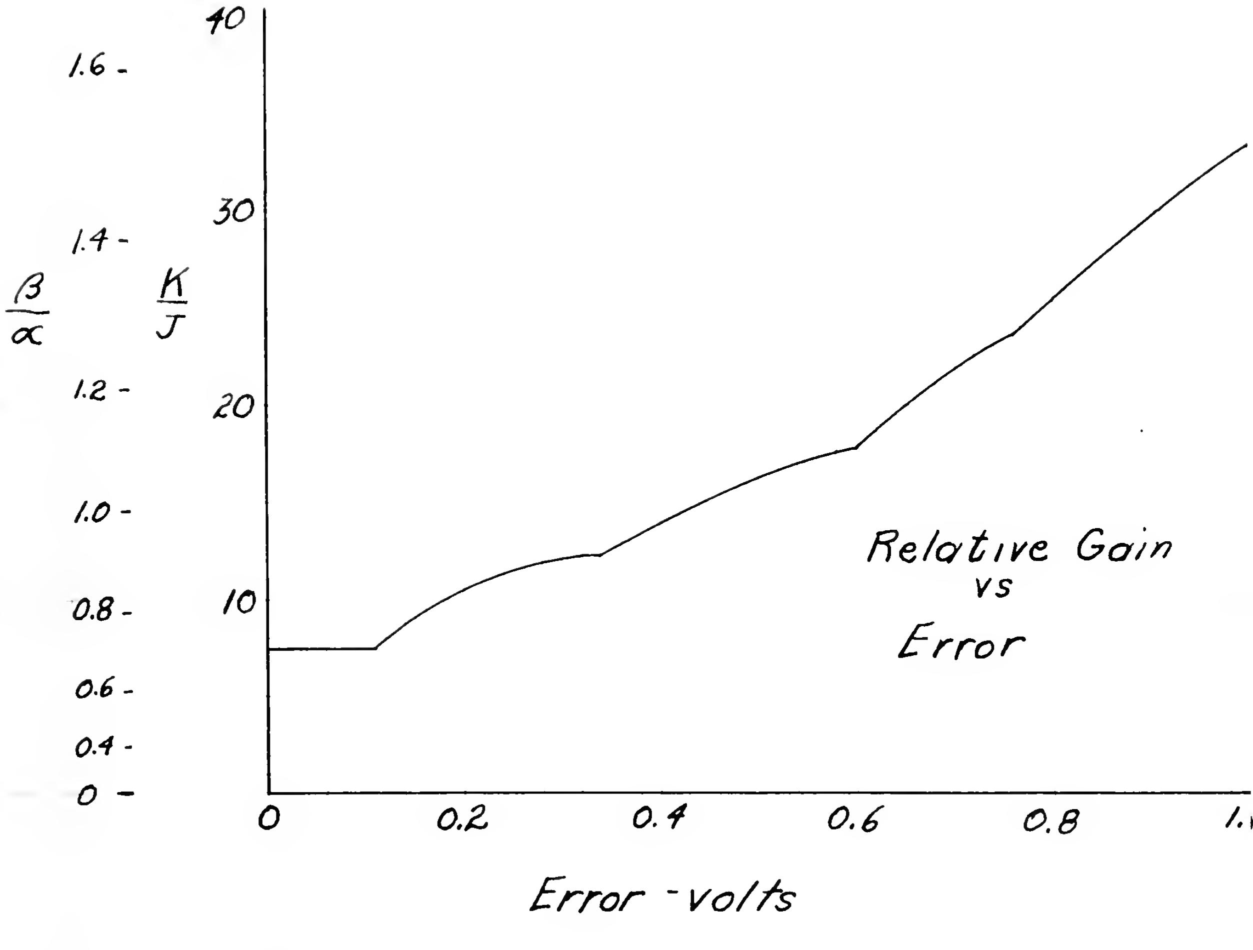


Figure XX

ANALYSIS OF DATA

The results of the tests made on the computer for the system operating with linear gain are in agreement with that predicted from the theoretical discussion earlier, with the exception that the optimum settling time of the system occurs when the gain is slightly in excess of that needed for a critically damped system. This can be explained by the fact that at this value of gain, the system is slightly underdamped but the overshoot at this gain is less than two percent, thus reducing the settling time to a value which is slightly less than the rise time of the system under the same conditions. The results of the system opera ing with linear gain were taken to be used as a basis of comparison for determining the effect of non-linear gain on the system.

When the system was operated with non-linear gain characteristics, it was found that the transient response of the system could be improved providing certain conditions were simultaneously met. These conditions will be set forth later.

In studying non-linear gain conditions, the first runs were made with the gain for zero error set at a value greater than that necessary for critical damping to verify or disprove the prediction that no appreciable improvement could be made with a non-linear gain characteristic such that the gain was at all times greater than that necessary for critical damping. This prediction was found to be true. Figures XXXII, XXXIII and XXXIV show some gain vs error characteristics and

the transient response obtained when using that characteristic for the system with the gain at zero error adjusted to be greater than that needed for critical damping in a linear system. There is no appreciable improvement in the settling time, and the overshoot and rise time are such that they correspond very nearly to that obtained when the linear system gain was adjusted to some mean value within the limits of the nonlinear characteristic. The overshoot and rise time of Figure XXXII is the same as that obtained on a linear system having $\frac{\beta}{\alpha}$ equal to 2.8, Figure XXXIII the same as $\frac{\beta}{\alpha}$ equal to 2.7 and Figure XXXIV the same as $\frac{\beta}{\alpha}$ equal to 2.0.

Another set of curves were taken with the gain vs error characteristic adjusted so that the gain for zero error was at the value needed for critical damping in a linear system. Figures XXXV, XXXVI and XXXVII show some of these gain vs error curves and the associated transient response. In all cases the system had overshoot and the settling time was about the same as could be obtained from a critically damped linear system. Actually this response was not as good as a linear system with critical damping because the linear system would settle to the final value in about the same time with no overshoot.

During the remainder of the investigation, the gain for zero error was set at a value below that needed for critical damping in a linear system. The gain vs error characteristic was then set to give increased gain for errors greater than

zero, and various shapes were used to study the effect on rise time, peak overshoot, and settling time. The effect of non-linear gain on all of these was found to be sensitive to the magnitude of the step function input. In many cases, the non-linear gain characteristic was found to have an adverse effect on the system response.

Figures XX and XXXVIII through XLII are all gain vs error characteristic which have the gain at zero error less than that needed in a linear system for critical damping. Figure XX has a characteristic which rises almost linearly with increased error. The transient response for this characteristic shows a rapid initial acceleration and then the system does damp out rapidly. The overshoot present is very small but it takes considerable time for this overshoot to decay because of the low gain at the error corresponding to this overshoot. This particular curve has a settling time very close to that of a linear, critically damped system, but has a faster rise time.

Figures XXXVIII, IXL and XL were all taken with the same value of gain at zero error, but the value of error at which the gain started to increase was changed. Figure XXXVIII has a characteristic which begins to increase at small errors and consequently the system has overshoot, caused by the fact that the system gain was too high during most of the transient period. Figure IXL has a gain vs error characteristic which does not start to rise until the error

signal is somewhat larger than it was in the previous figure. The system still has some overshoot. The next figure has a characteristic that is high during only a small part of the transient period, and has no overshoot and a very good settling time of only 3.0 time constants. This is the best achieved so far.

Figure XLI shows the effect of changing the amplitude of the input signal. The zero error gain is lower than that used before, but the increase in gain for increased error is more abrupt. The settling time, in general, is the same as that for an overdamped system having the same linear gain as this system has for zero error. When the input signal was small, very little effect of the inc easing gain characteristic was noticed. When the input signal was large, the system operated at high gain during too much of the transient period and had considerable overshoot corresponding to this higher average gain. When the input was 1.0 volts, the correct combination of high and low gain was achieved, and the system settled to the final value in a very short time, 2.2 time constants. Figure XLII also shows the effect of changing signal amplitude on a fixed gain vs error characteristic. The gain for zero error is lower than any previous curve. Due to this lower gain, the system had a very long settling time in general, but at the value of input (1.0 volts) which gave the right combination of high and low gain, the performance during the transient period was very good.

A characteristic common to all of the gain curves which produced improvements in performance is the rapid rise in the gain characteristic, and that the gain at zero error was much smaller than that needed for critical damping. The value of error at which the gain starts to increase abruptly is dependent upon the value of the input signal if optimum transient performance is to be obtained. This led to the concept of using a gain vs error characteristic which was very low for errors below some finite value, and then increased with infinite slope to some high value for error larger than this value. This concept was studied theoretical and experimentally and the results are detailed in later paragraphs, following some general discussion defining the concept of improved response.

The concept of improved response as produced by the non-linear amplifier may be considered from two points of view, that of faster rise due to the high gain applied initially, and that of decreased settling time due to increased damping (decreased gain) as the error reduces. For certain amplitudes of step input, the rise time seems to be the greatest improvement, but this is misleading, because small step inputs result in a greatly increased rise time. For a linear system, a fast rise time is an indication of the steady state error and velocity error of the system, and it was pointed out earlier that the low gain and long rise time required for good damping would result in large steady state errors and velocity

lag errors, thus rise time considerations are misleading in judging the improved transient response of a non-linear system, since the rise time depends on the magnitude of the disturbance. The best way to obtain a consistently short rise time is to use a linear high gain but this results in an underdamped system, having considerable overshoot and oscillations before settling to the final value.

On the other hand, if settling time is used as a measure of improved performance, certain interesting results are obtained. In general, the settling time of a system using non-linear gain is longer than for an underdamped linear system, approaching as an upper limit the settling time of a system having a linear gain equal to the zero error gain of the non-linear system. This upper limit on the settling time is approached for both large and small values of step displacement inputs. For some intermediate value of step input, the settling time decreases to a minimum of two time constants, which is considerably better than can be obtained with a linear system of the same complexity. These conditions are shown to exist in Figure XLII-d.

It is therefore obvious that system response, in terms of settling time, can be improved by a non-linear gain characteristic providing the proper combinations of conditions can be obtained. The following paragraphs develop the mathematical theory leading to a definition of the conditions necessary for optimum settling time with non-linear gain.

The experimental results led to the conclusion that the optimum system performance could be achieved when the gain vs error characteristic for the non-linear system was adjusted to give a very low but finite value of gain for errors between zero and some finite value, then increasing abruptly for errors larger than this value. To facilitate mathematical study of this characteristic, a gain vs error characteristic which had a gain of K_2 for errors less than some value δ , and which increased with infinite slope to some value K_1 for errors larger than δ was used. This characteristic is shown in Figure XXI. The effect of allowing K_2 to approach zero as a limit was studied, and the effect of allowing K_1 to approach infinity was also studied.

The system having the gain characteristic shown in Figure XXI was assumed to be at rest and to have a step displacement input of Θ at time equal to zero. It was also assumed that K_1 was of such large magnitude that the linear system with this amount of gain would be underdamped. That is:

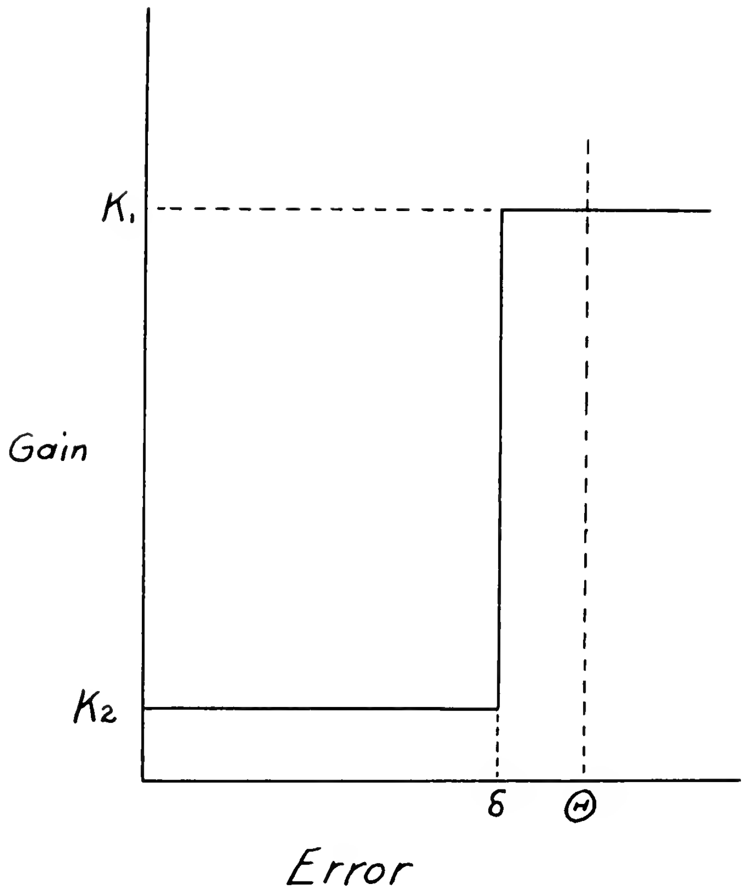
$$\frac{K_1}{J} > \left(\frac{F}{2J}\right)^2$$

The differential equation of the system under this condition is:

$$J\frac{d^2\theta_o}{dt^2} + F\frac{d\theta_o}{dt} + K_1\theta_o = K_1\Theta \qquad (15)$$

and has the solution:

$$\theta_o = \Theta\left[1 - \varepsilon^{-\alpha t}\left(\frac{\alpha}{\omega_1}\sin\omega_1 t + \cos\omega_1 t\right)\right] \qquad (16)$$



Theoretical Gain vs Error Characteristic

Figure XXI

where: $\alpha = \dfrac{F}{2J}$ $\beta_1^2 = \dfrac{K_1}{J}$

$$\omega_1^2 = \beta_1^2 - \alpha^2 = \frac{K_1}{J} - \left(\frac{F}{2J}\right)^2$$

The expression for the output velocity at any time is given by:

$$\frac{d\theta_o}{dt} = \Theta\, \epsilon^{-\alpha t} \frac{\beta_1^2}{\omega_1} \sin \omega_1 t \qquad (18)$$

It is now necessary to find the value of $\frac{d\theta_o}{dt}$ when the error has reduced to δ , so that this may be used as an initial condition for the system response when the gain has reduced to K_2. This value of the velocity will be designated Ω_o .

This may be done by solving the equation for:

$$\theta_o = \Theta - \delta = \Theta \left[1 - \epsilon^{-\alpha t} \left(\frac{\alpha}{\omega_1} \sin \omega_1 t + \cos \omega_1 t \right) \right] \qquad (19)$$

or:

$$\delta = \Theta\, \epsilon^{-\alpha t} \left(\frac{\alpha}{\omega_1} \sin \omega_1 t + \cos \omega_1 t \right) \qquad (20)$$

to find the time, then substituting this to find the value of $\frac{d\theta_o}{dt}$ at this time. This cannot be expressed analytically because the equation for θ_o does not have a formal solution. It must be solved by some method such as a graphical solution or by successive approximations.

If, however, we assume K_1 is so large that the velocity term of the differential equation is negligible compared to the acceleration term, we can get an approximate equation of:

$$\theta_o \doteq \frac{\Theta \beta_1^2 t^2}{2} \qquad (21)$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

and:

$$\frac{d\theta_o}{dt} \doteq \Theta\beta_1^2 t \tag{22}$$

for small values of t. Hence, the velocity of the system when the error has reduced to δ can be approximated by

$$\Omega_o \doteq \beta_1\sqrt{2\Theta(\Theta-\delta)} \tag{23}$$

and the time to reach this position is given by:

$$t \doteq \frac{1}{\beta_1}\sqrt{\frac{2(\Theta-\delta)}{\Theta}} \tag{24}$$

to find the equation of motion for the system after the gain has reduced to K_2, it is most convenient to shift coordinates such that the system is now considered to have a step displacement input of δ , and the system has an initial velocity Ω_o at time equal zero. It will be assumed that the system gain K_2 is below that needed in a linear system for critical damping.

The equation of motion is then:

$$\theta_o = \delta\left[1-\epsilon^{-\alpha t}\left(\frac{\alpha}{\gamma_2}\sinh\gamma_2 t + \cosh\gamma_2 t\right)\right] + \frac{\Omega_o}{\gamma_2}\epsilon^{-\alpha t}\sinh\gamma_2 t \tag{25}$$

where

$$\gamma_2^2 = \alpha^2 - \beta_2^2 = \left(\frac{F}{2J}\right)^2 - \frac{K_2}{J}$$

This can also be expressed as:

$$\theta_o = \delta\left[1-\epsilon^{-\alpha t}\left(\frac{\alpha-\frac{\Omega_o}{\delta}}{\gamma_2}\sinh\gamma_2 t + \cosh\gamma_2 t\right)\right]$$

The curves of this equation are very similar to the curves

obtained in Figure XLII. It can be shown that this equation may give an overshoot, even though not oscillatory in nature, depending on the relative magnitude of Ω_o.

If K_2 is allowed to approach zero as a limit, then θ_o is given by:

$$\theta_o = \frac{\Omega_o}{2\alpha}\left[1 - \epsilon^{-2\alpha t}\right] \tag{27}$$

which settles to $\frac{\Omega_o}{2\alpha}$ in two time constants. For the system to come to correspondence, it is necessary that:

$$\theta_o = \frac{\Omega_o}{2\alpha} = \delta \tag{28}$$

Thus, the time during which the system is at high gain must be such that when the error has reduced from Θ to δ , the velocity, Ω_o , is equal to $2\alpha\delta$. With a fixed δ and a given value of Θ , K_1 must be adjusted to meet this condition, if K_1 is fixed and with a given value of Θ , δ must be adjusted, or if K_1 and δ are fixed, Θ must be of the correct magnitude or the system will not come to rest at the desired position.

If K_2 is very small, the effect of K_2 is not appreciable on the position which the system will tend to approach, namely:

$$\theta_o = \Theta + \frac{\Omega_o}{2\alpha} - \delta \tag{29}$$

but will effect how long it takes the system to change from this position to the desired position.

If K_1 approaches ∞ , then for a given value of Θ , if

is to remain finite, ($\Theta - \delta$) must approach zero, which is equivalent to an infinite torque operating through zero distance. This impulse must be such that

$$\Omega_0 = 2\alpha\Theta \tag{30}$$

(if K_2 is zero) if the system is to come to rest at the desired final value.

When the characteristic is adjusted to have K_1 large, but finite, and K_2 is set to zero, the system operates essentially as a relay servo. The value of δ is indicative of the "dead zone" of the relay servo.

If K_2 is not zero, but is small, it results in a relay type of servo that has the advantage over a conventional relay servo of having position control in the "dead zone", but operation in this "dead zone" is very sluggish due to the low gain (high damping).

As K_2 is allowed to approach K_1, the system approaches linear operation. The effect of the non-linear gain is to give a system which has characteristics common to both the linear servo and the relay servo.

The fact that the shape of the gain vs error curve must be different for each different value of step disturbance applied to the system leads to some interesting developments

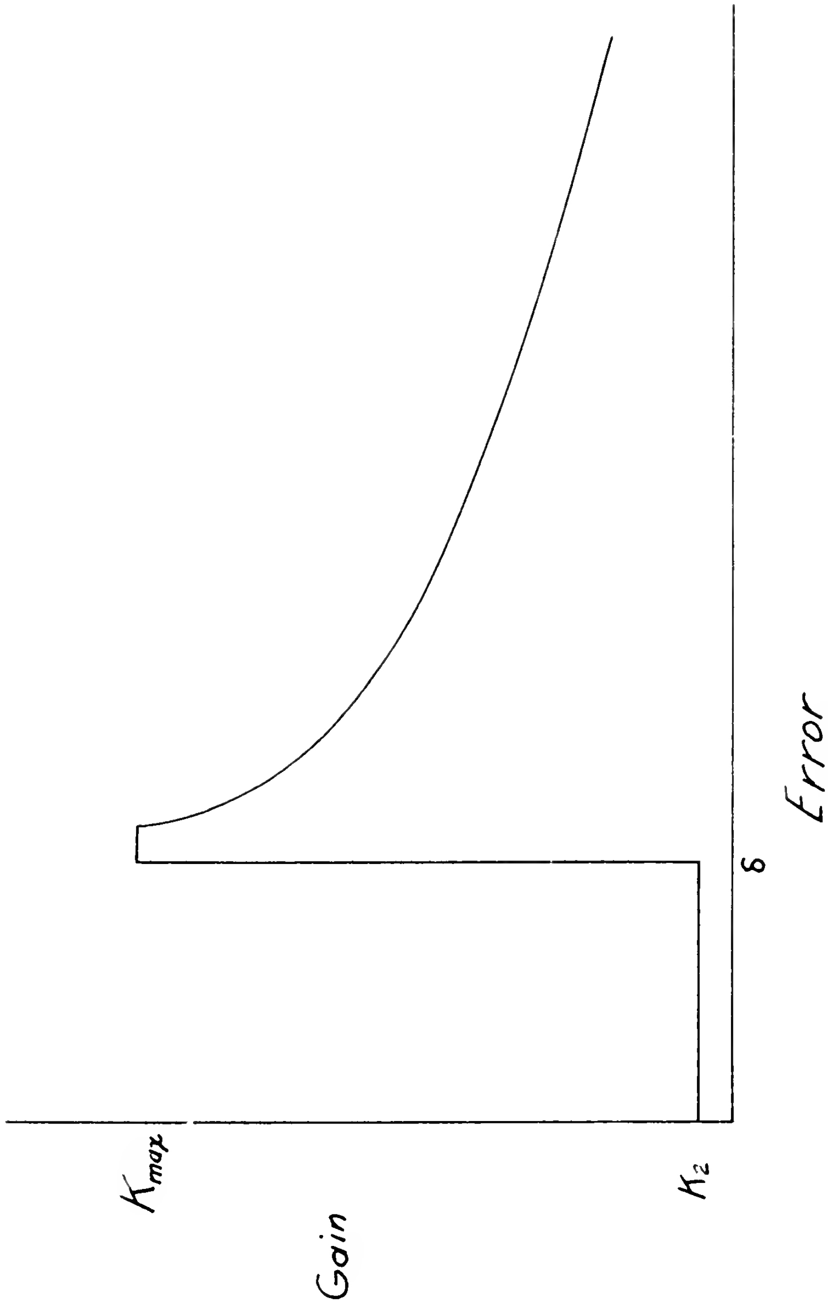
for possible use of non-linear gain characteristics to achieve a system which has a settling time of about two time constants for all values of step disturbance, or at least over a range of values, rather than being restricted to only one value as was the case in this investigation.

The factor that is of prime importance in obtaining correspondence in a minimum time is that the velocity must be the correct magnitude when the gain is reduced in order that the system will come to rest at the desired value. The amount of the velocity, for a step input, was shown earlier to be:

$$\Omega_0 = 2\alpha\delta \tag{31}$$

if K_2 is zero. If K_2 is not zero, the value of velocity to make the system approach the desired value in the shortest time may be found from a solution of equation (26), or by experimentation. This leads one to believe that there may be a possible gain vs error characteristic which would have a very low value of gain for errors less than a fixed value, δ, and which has a gain vs error curve for errors larger than δ which would produce the desired value of Ω_0. At the instant that the error had been reduced to δ, independent of the magnitude of the step disturbance provided only that this magnitude was greater than δ. A gain vs error curve of this might be similar to the one shown in Figure XXII.

It is felt, however, that this approach to the solution of giving an improved performance over a wide range of step



A Proposed Gain vs Error Characteristic

Figure XXII

displacements would not prove practical because of two factors. The first being that with a fixed value of δ, the system response would be very sluggish for disturbances which had an amplitude less than δ. If δ is made small to reduce this sluggish region, the velocity which the system is to attain at the instant the error had been reduced to δ would necessarily be relatively small, resulting in a relatively long time for the system output to change from the value of the disturbance to the value δ, in addition to the two time constants required to come to rest after the error had been reduced to δ.

Another method of control which probably has much greater possibilities in attaining a system with improved performance is one in which the value of K_1 is fixed at some very large value, and have the value of δ vary to control the distance which would be required to bring the system to rest in the shortest time at the desired value after the gain had been reduced to zero. This would result in a system which would attain a relatively high velocity as soon as the disturbance was applied, because of the high gain initially applied, and which would result in a system which had no "dead zone" and which would give a settling time of about two time constants, with no overshoot, for any value of disturbance. The time by which the settling time differed from two time constants would be the time it took the system to reach the velocity, Ω_o, which would be a function of the value of the

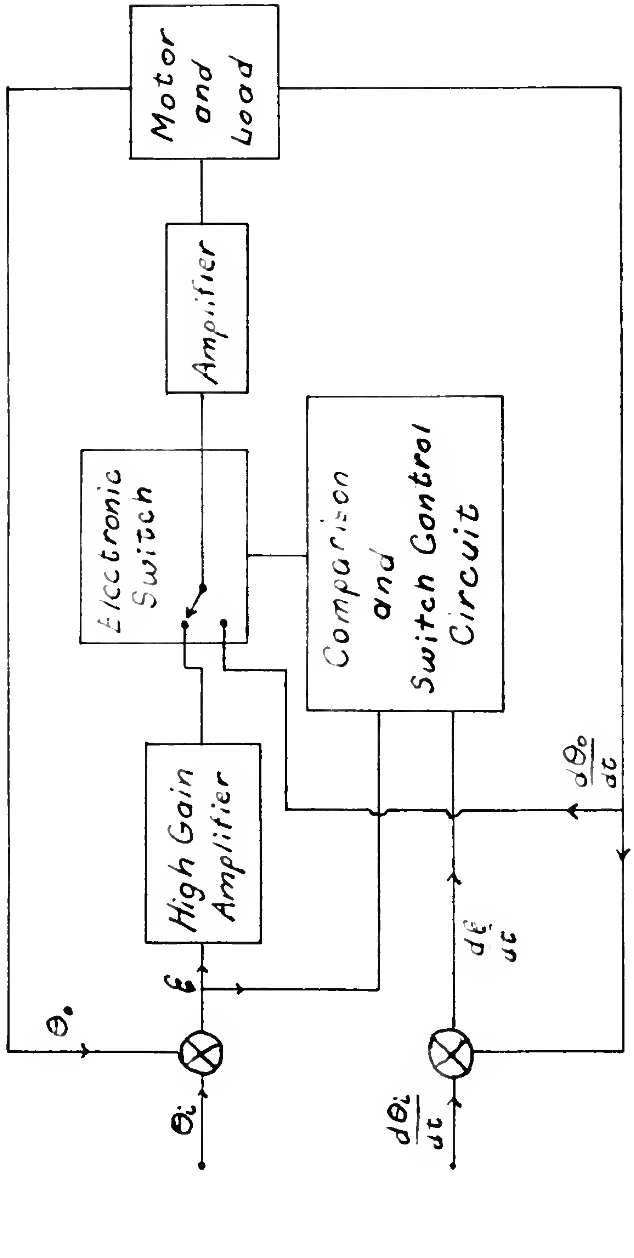
disturbance and which would be relatively small due to the very high value of gain used when the disturbance was first applied.

This method of control would be such that the velocity of the system would be compared against the error of the system, and if the velocity was less than $2\alpha\mathcal{E}$, the system would have high gain. When the velocity reached a value of $2\alpha\mathcal{E}$, the system would turn off. If at the time that the high gain was turned off, a velocity feedback was turned on, the time constant of the system would be very greatly reduced. It is not desired to have the increased damping when the high gain is on because it would increase the acceleration time of the system. If anything, the velocity feedback would be reversed during the time the high gain was on to reduce the effect of the viscuous friction during the acceleration time.

A block diagram of this proposed system is shown in Figure XXIII, and a phase plane diagram of the system operating characteristics is shown in Figure XXIV. The operation of the system is such that when the disturbance is first applied, the high gain channel is turned on. When the velocity has reached a value such that:

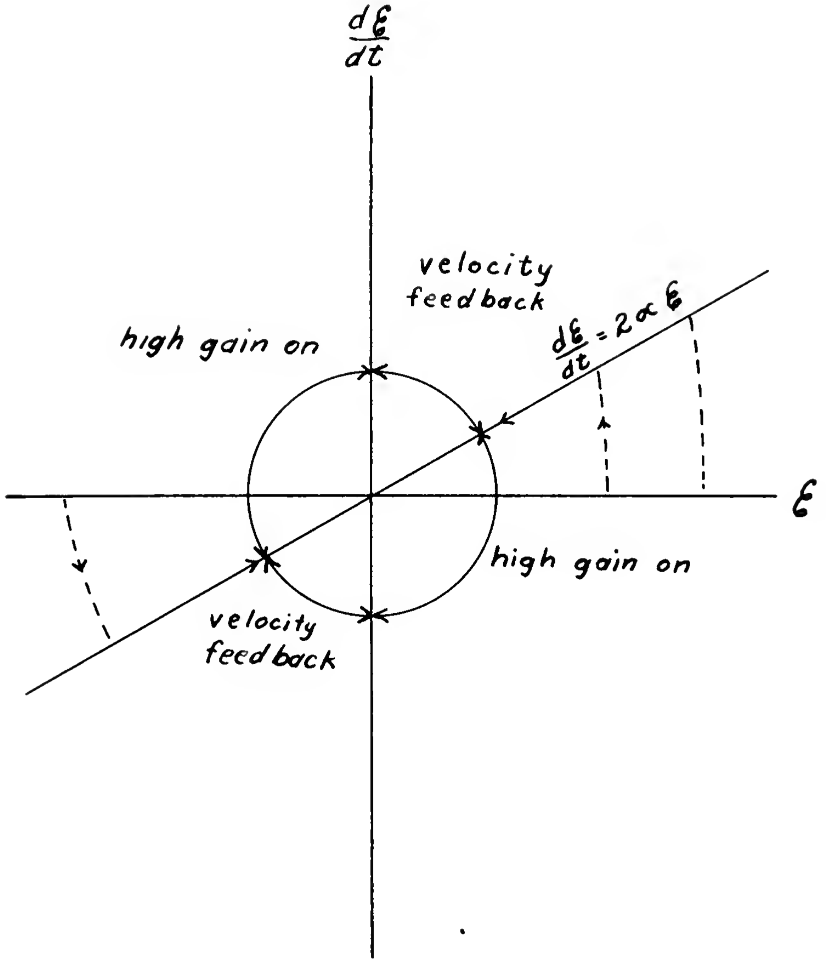
$$\frac{d\theta_o}{dt} = 2\alpha\mathcal{E} \qquad (32)$$

the high gain channel is turned off and the velocity feedback channel is turned on. The switching would preferably be accomplished by an electronic switching circuit to avoid the



Block Diagram of a Proposed Improved System

Figure XXIII



Phase Plane Diagram

Figure XXIV

time delays and mechanical difficulties encountered in mechanical relays. The exact details of the circuit of the comparing and switching device has not been developed, but it is not felt that this would be too difficult or complicated.

For disturbances which are not step functions, the rate of change of error should more properly be used to determine the time at which the system was turned off, rather than the output velocity. This could be accomplished in practice very easily by installing tachometers on both the input and output position shafts, and taking the difference of thier outputs as the measure of the rate of change of error, providing the addition of the tachometer to the input shaft would not overload the driving source.

CONCLUSIONS

It has been concluded that non-linear gain, which has a characteristic such that the gain increases with increased system error, can produce a marked effect in the transient performance of the system. The non-linear gain characteristic can reduce the settling time from a minimum of about four time constants, or six time constants if the system is not to have overshoot, in a linear system to an optimum of two time constants, without overshoot, in the non-linear system, providing certain conditions can be met.

These conditions are: 1) that the gain at zero error, K_2, must be well below that needed in a linear system for critically damping; 2) the gain, K_1, at errors larger than some value of error, δ , must be very high and 3) the step disturbance must be the correct value for the values of K_1, K_2, and δ chosen.

If these conditions are not met, the effect of non-linear gain is, in general, detrimental on the system response. The resulting effect of the non-linear gain is to give a system which has a response that combines the on-off characteristics of a relay servo with the accurate positioning (elimination of the "dead zone") of a linear system.

It has been proven, theoretically and experimentally, that a gain characteristic which gives increased gain for increased error is of no practical value. The investigations

indicate that other types of non-linear gain variation may lead to a system which has improved performance over a linear system.

RECOMMENDATIONS FOR FURTHER INVESTIGATION

It is felt that the method of controlling the value of error at which the high gain of the system was reduced to a very low, or zero, value offers a good project for further investigation of non-linear gain characteristics and their effect on the transient response of a servomechanism system. A more detailed description of this method was given on pages 62 to 64 and further discussion will not be repeated here.

Another method which shows some promise, but not as much as that mentioned above, is the shaping of the gain characteristic curve to give the same value of velocity at the error value at which the gain is reduced to a low value, regardless of the amplitude of the disturbance as long as the disturbance was larger than the minimum error for high gain. This was also discussed earlier on pages 60 to 62..

The introduction of a derivative damper in connection with a fixed non-linear gain characteristic, to operate on the error signal either before or after the non-linear amplifier, was thought to be able to produce some effect on the transient behavior of the system, but whether it would be beneficial is not known. Further investigations along this line might also be warranted.

BIBLIOGRAPHY

1. Brown, G.S. and Campbell, D.P.: Principles of Servo-mechanisms

2. Gardner, M.F. and Barnes, J.L.: Transients in Linear Systems - Volume I

3. Ginzer: "D-C Amplifier Design Techniques". Electronics, Vol. 17, 1944, page 98.

4. Stout, T.M.: A Study of Some Nearly Optimum Servo-mechanisms. Electrical Engineering Report No. 10 of November 28, 1952; University of Washington

5. Thaler, G.J. and Brown, R.G.: Servomechanism Analysis

APPENDIX A

DESIGN AND CONSTRUCTION OF THE NON-LINEAR AMPLIFIER

The diode limiter type of non-linear circuit was chosen to provide the gain variation with signal amplitude because of the wide range of operating characteristics permitted by the adjustment of the circuit parameters. The circuit used is shown in Figure IX.

The center tap of the bias supply is tied to the output terminal of the circuit, therefore a type of supply had to be selected that would float with the output and not introduce an effective load on the circuit. A battery seemed to be the logical choice for this investigation, since it has no connection to any outside power supply or ground, but an isolated electronic power supply could be used. The use of the direct current supply to the laboratory is definitely not permissible due to circuit interaction.

The choice of the value of resistors to use in the circuit was determined by two restrictions; 1) the potentiometers used to adjust the bias voltage to the diodes had to be sufficiently large that they did not put too much drain on the battery; 2) these same potentiometers had to be small compared to the values used in the remainder of the circuit so that changing the bias voltage to the diodes did not upset the calibration of the resistors used in series with the

diodes. The latter restriction was imposed to permit calibration of the variable resistors, so that the circuit characteristics could be computed, rather than run performance curves for each adjustment. The value for each potentiometer was chosen as ten thousand ohms. This offered a resulting load to the battery of five thousand ohms, and gave a drain of forty milliamperes of the battery if a 20 volt source was used. In the investigation, a 22½ volt battery was used with a variable series resistor to adjust the value at the potentiometers to 20 volts. This permitted calibrating the potentiometers directly in volts, and being able to duplicate the settings readily.

Since the resistance of the potentiometers was to be small in comparison to the resistance of the other circuit parameters, the value of Ro was chosen to be 100 thousand ohms. After construction was completed and the values measured on a bridge, this resistor was found to have a value of 81.2 thousand ohms.

It had been decided that a change in the gain of the circuit should be on the order of at least five or six to one. This lead to a choice for R of about ten times Ro, or one million ohms. Measurement of the resistor used showed it to have a value of 1.04 megohms.

The values of the variable resistors in series with the diodes had to be such that the parallel resistance of these with R could be changed uniformly down to a very small value.

The net result of the diodes and variable resistors is that it gives a circuit which has a signal applied to a fixed resistor and a variable resistance in series, with the output being the value of voltage across the fixed resistor. The magnitude of the variable resistor is a function of the applied voltage.

The operation of the non-linear circuit is typified by the curve of output vs input shown in Figure XXV. In the region between O and A on the curve, all diodes are non-conducting because of the bias applied to them. The slope of the curve between these two points is then equal to:

$$\frac{R_o}{R + R_o} \tag{33}$$

where Ro includes the effect of the input impedance of the following stage. The output impedance of the preceding stage should also be taken into account, but it will be assumed that this is negligible. This assumption will be shown to be true later.

Then the voltage across R, or the difference between the input and the output voltages, increases until it is equal to E_A, the first diode begins to conduct, reducing the resistance R to a value which is the parallel resistance of R and R_A. Thus, the slope of the section of the curve between A and B is:

$$\frac{R_o}{\frac{R \cdot R_A}{R + R_A} + R_o} \tag{34}$$

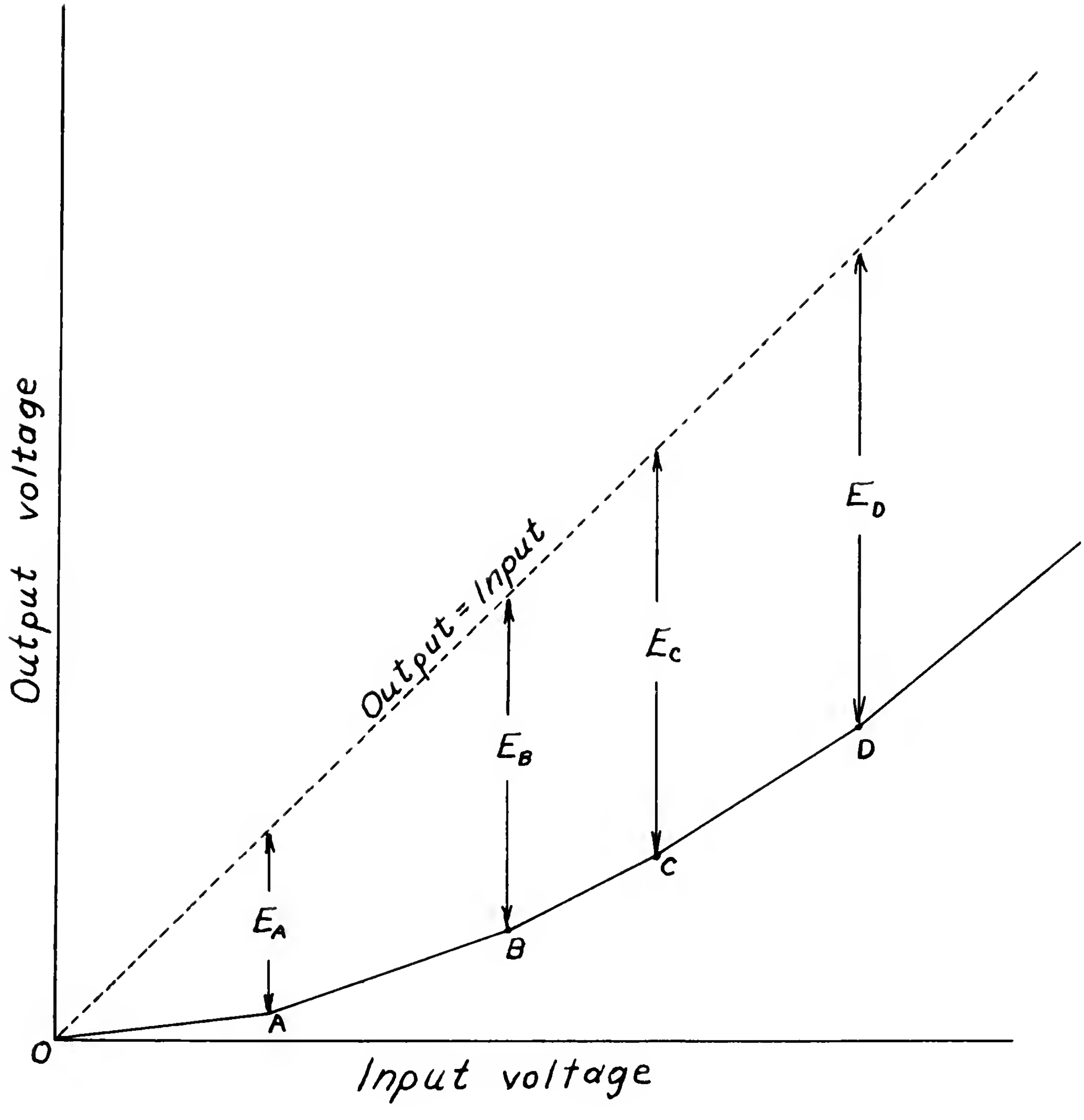


Figure XXV

Similarly, when the voltage across R is between E_B and E_C, the slope of the curve is:

$$\frac{R_o}{\dfrac{1}{\frac{1}{R}+\frac{1}{R_A}+\frac{1}{R_B}}+R_o} \qquad (35)$$

Between C and D the slope is:

$$\frac{R_o}{\dfrac{1}{\frac{1}{R}+\frac{1}{R_A}+\frac{1}{R_B}+\frac{1}{R_C}}+R_o} \qquad (36)$$

and beyond D, the slope is:

$$\frac{R_o}{\dfrac{1}{\frac{1}{R}+\frac{1}{R_A}+\frac{1}{R_B}+\frac{1}{R_C}+\frac{1}{R_D}}+R_o} \qquad (37)$$

The gain for this circuit is then found by dividing the output voltage by the corresponding value of input voltage. The output vs input curve and the gain vs input curve for the circuit, with the circuit values as listed below, are shown in Figures XXVI and XXVII respectively. The input impedance of the stage following the non-linear circuit was taken to be one megohm, giving an effective Ro of 75 thousand ohms.

Circuit Values

R_A	1.0 Megohm	E_A	1 volt
R_B	0.5 Megohm	E_B	3 volts
R_C	0.2 Megohm	E_C	5 volts
R_D	0.1 Megohm	E_D	8 volts

The overall system gain is found by multiplying the gain of all circuits exclusive of the non-linear circuit by the gain of the non-linear circuit. The gain of the other circuits

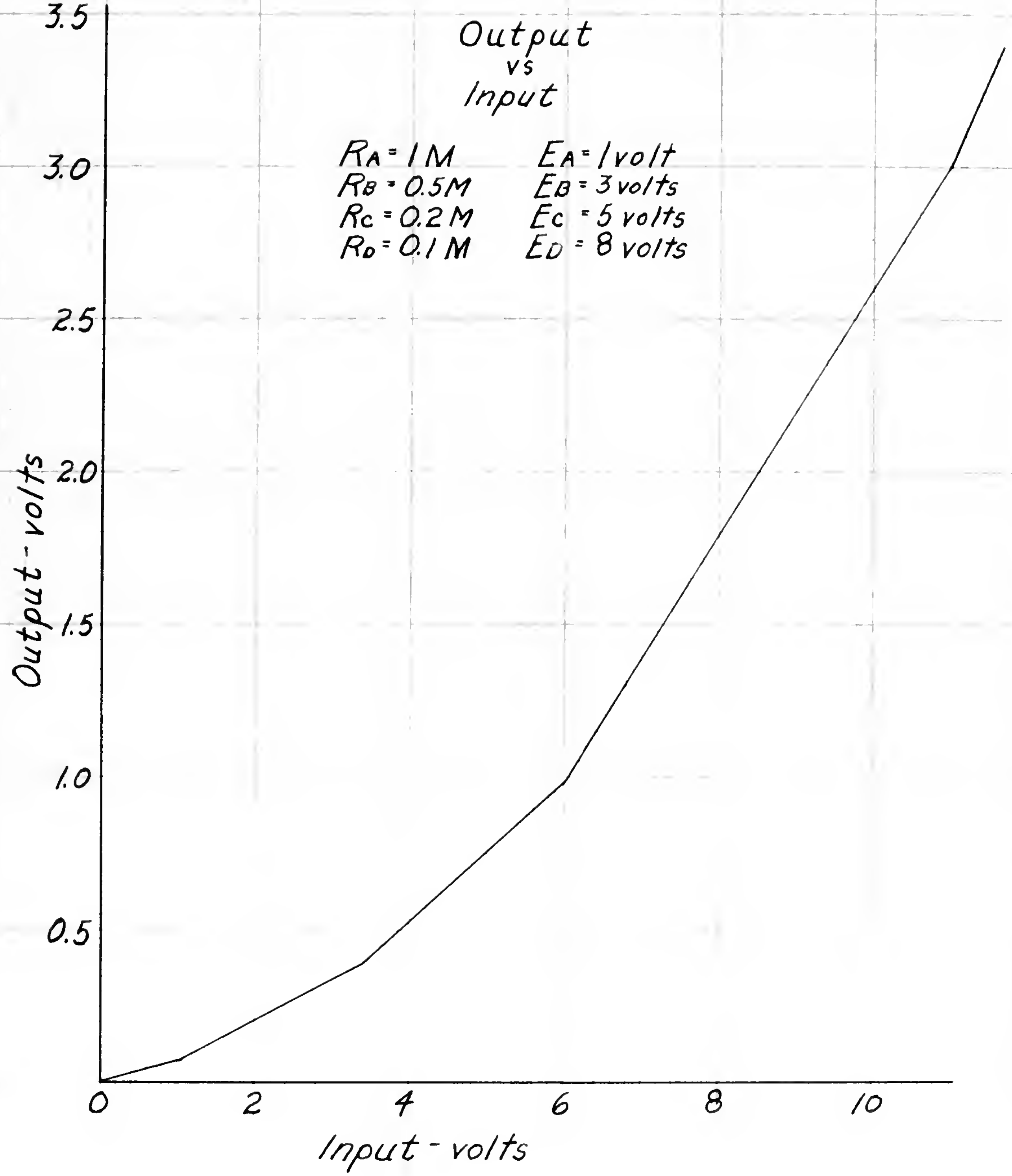

Output
vs
Input
RA = 1 M
RB = 0.5M
RC = 0.2 M
RD = 0.1 M
EA = 1 volt
EB = 3 volts
EC = 5 volts
ED = 8 volts
3.5
3.0
2.5
2.0
1.5
1.0
0.5
Output - volts
0
2
4
6
8
10
Input - volts

Figure XXVI

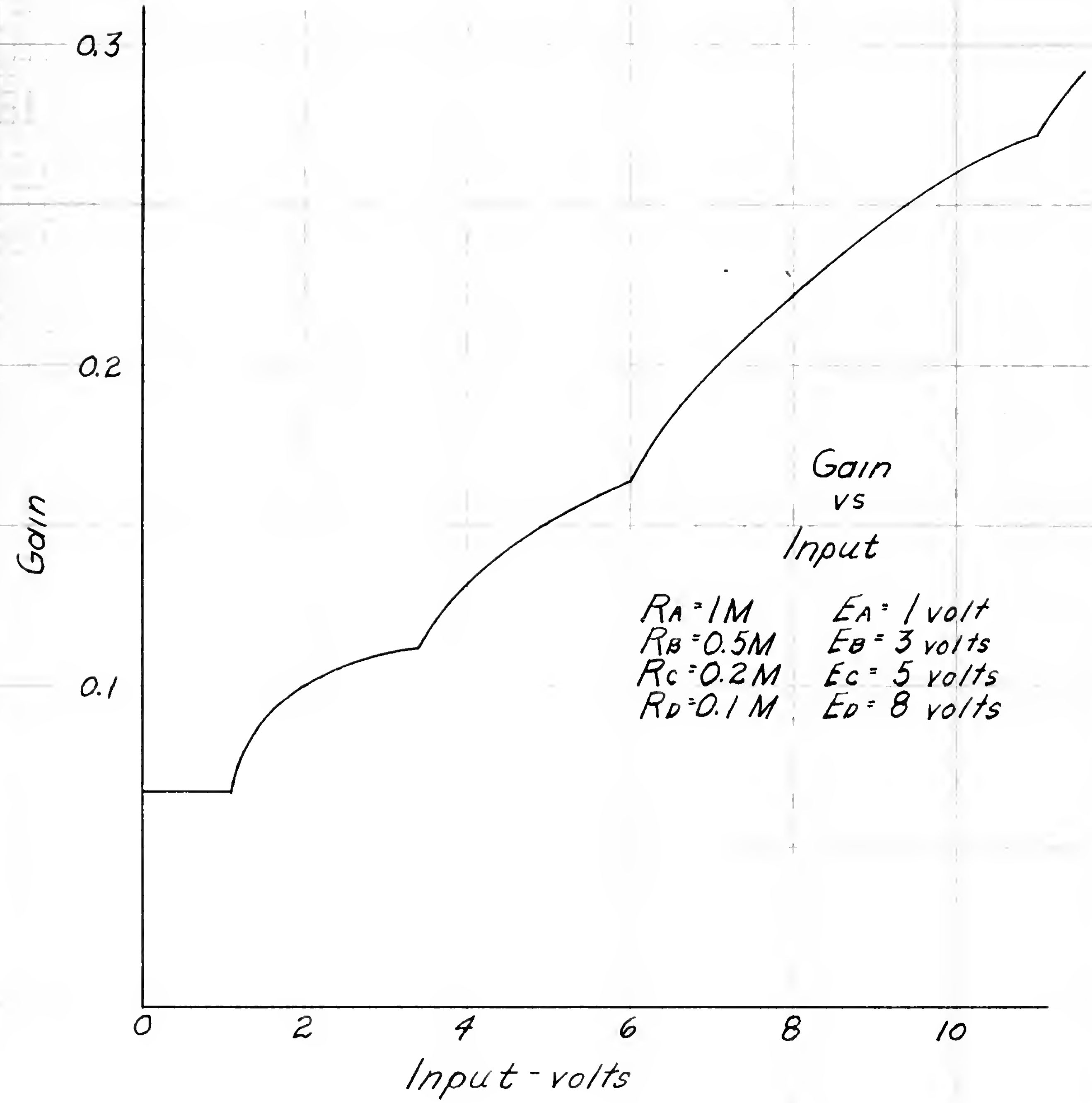


Figure XXVII

are linear within the range of voltages used, and are represented by a constant numerical value.

At the time that the non-linear circuit was built, it was considered necessary that there be a preamplifier between the non-linear circuit and the error detector. The error detector was to be d.c., therefore a d.c. amplifier was designed and incorporated into the same chassis. The requirements of the preamplifier were that it have an overall gain of about ten, that it present a negligible source impedance to the non-linear circuit, and that it have minimum drift. A Miller circuit was used as the input stage of the d.c. amplifier to minimize the drift. This stage was found to have a gain of about three and one half. An additional stage of gain was added to meet the gain requirements and a cathode follower was added as the last stage to give a low output impedance for the amplifier. The d.c. amplifier circuit is shown in Figure VIII.

If the system had been originally designed to use an a.c. error detector, the preamplifier could more easily have been an a.c. amplifier added before the chopper circuit, instead of using the d.c. amplifier that was actually used. This would have done away with the more inherent difficulties found in d.c. amplifiers, namely, direct coupling and drift. Since the d.c. amplifier was built before it was found necessary to introduce the synchro error detector, it was left in the circuit when the conversion was made, rather than build an additional a.c. amplifier.

[illegible]
[illegible]

[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

APPENDIX B

RESULTS OF TESTS

The information contained in this appendix is the results of the tests made on the analog computer, for the circuit shown in Figure XVII. Figures XXVII through XXXI are the results of tests with linear gain employed, and the information is included as a basis of comparison for the effect of non-linear gain. The discontinuities in the settling time vs gain curve in Figure XXXI are due to the manner in which the settling time was defined. The first section of the curve is for the system in either the overdamped condition, or very slightly underdamped with the peak overshoot less than two percent. At the value of gain for which the first overshoot exceeds two percent, there is an abrupt increase in settling time. Similarly, as the number of overshoots and undershoots increases, there will be a discontinuity in the settling time, the amount of such discontinuity decreasing as the number of overshoots and undershoots increases.

Figures XXXII through XLI show the curves of gain vs error and the transient response of the system under such gain characteristics.

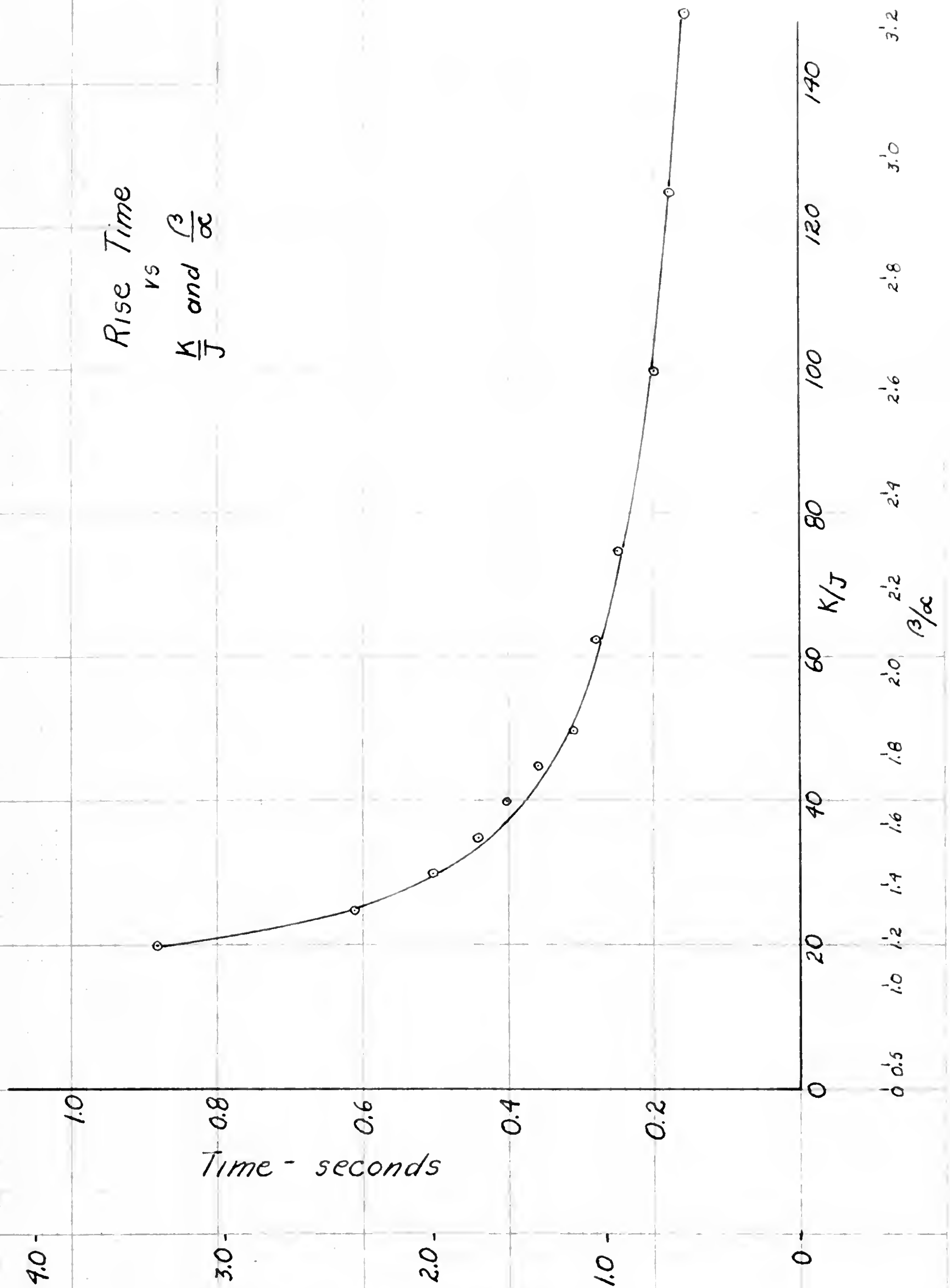
TABLE OF RISE TIME, SETTLING TIME AND PERCENT OVERSHOOT VS GAIN FOR A LINEAR SYSTEM

$\frac{K}{J}$	$\frac{\beta}{\alpha}$	Rise Time	Settling Time	Percent Overshoot
5	0.59	-	17.6	-
10	0.82	-	7.8	-
15	1.02	-	4.5	-
20	1.18	3.3	3.0	2
25	1.32	2.3	4.3	5
30	1.44	1.9	4.0	8
35	1.56	1.7	3.9	10
40	1.66	1.5	3.5	12
45	1.77	1.4	4.2	14
50	1.86	1.2	4.2	16
$62\frac{1}{2}$	2.08	1.1	4.0	20
75	2.28	1.0	3.6	25
100	2.63	0.8	4.2	30
125	2.94	0.7	3.8	35
150	3.22	0.6	4.1	38

System $\frac{F}{J} = 7.6$ System time constant, 0.26 sec.

Rise time and settling time in time constants

Figure XXVIII



System Time Constants

Figure XXIX

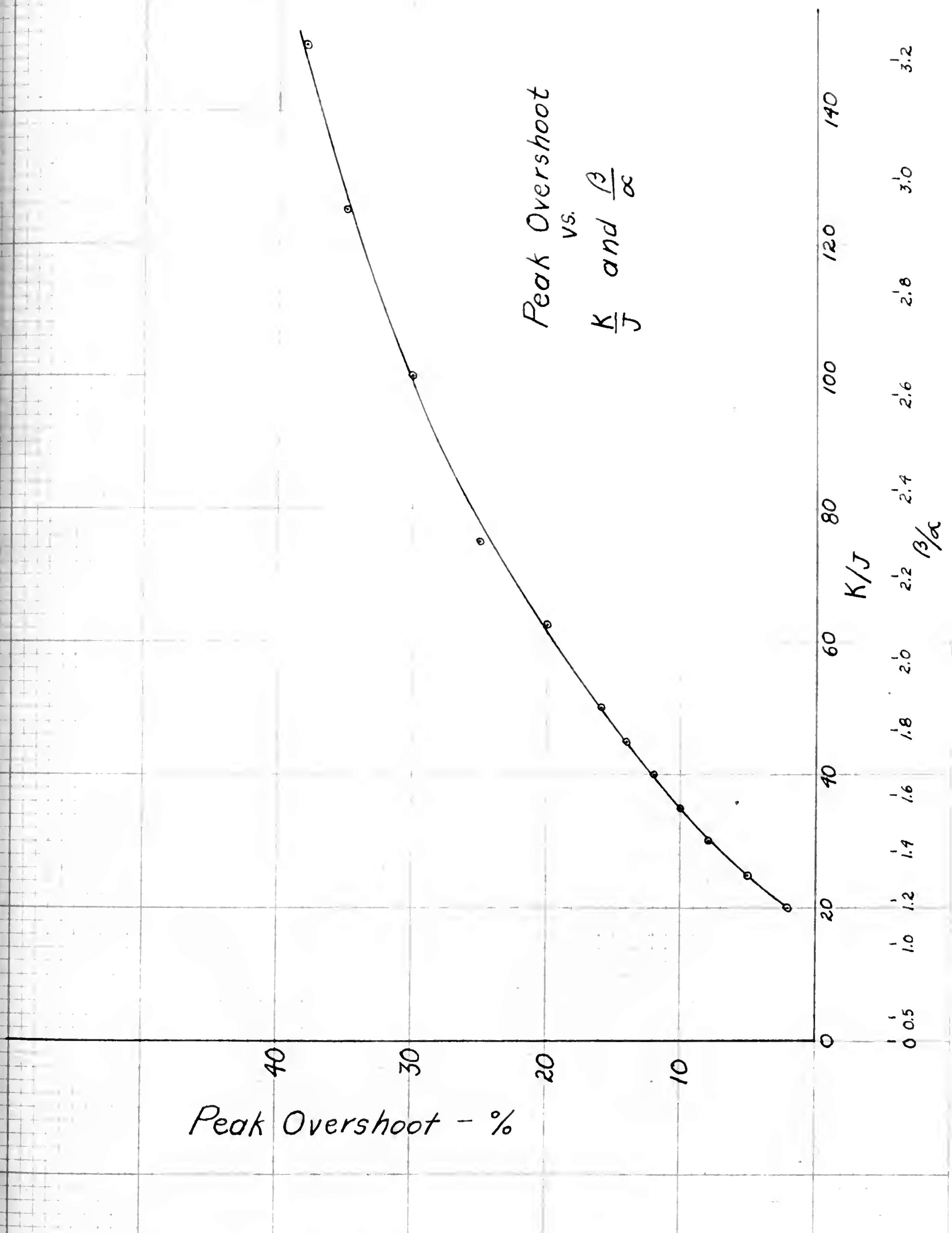

Peak Overshoot vs. K/J and β/α
Peak Overshoot - %
40
30
20
10
K/J
0
20
40
60
80
100
120
140
β/α
0 0.5
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2

Figure XXX

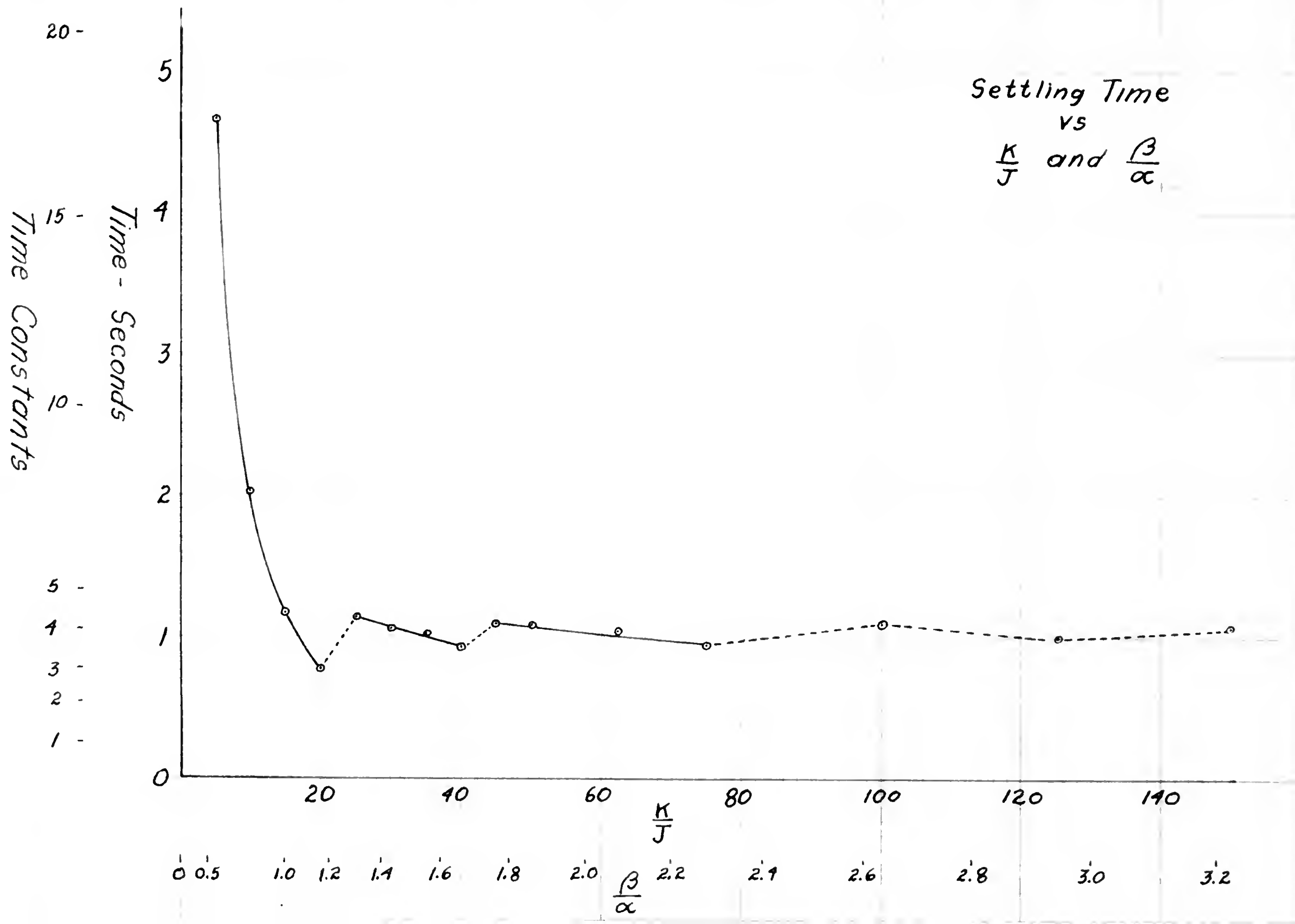

Settling Time
vs
K/J and β/α
Time - Seconds
Time Constants
K/J
β/α
0
1
2
3
4
5
20
40
60
80
100
120
140
0
0.5
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
1
2
3
4
5
10
15
20

Figure XXXI

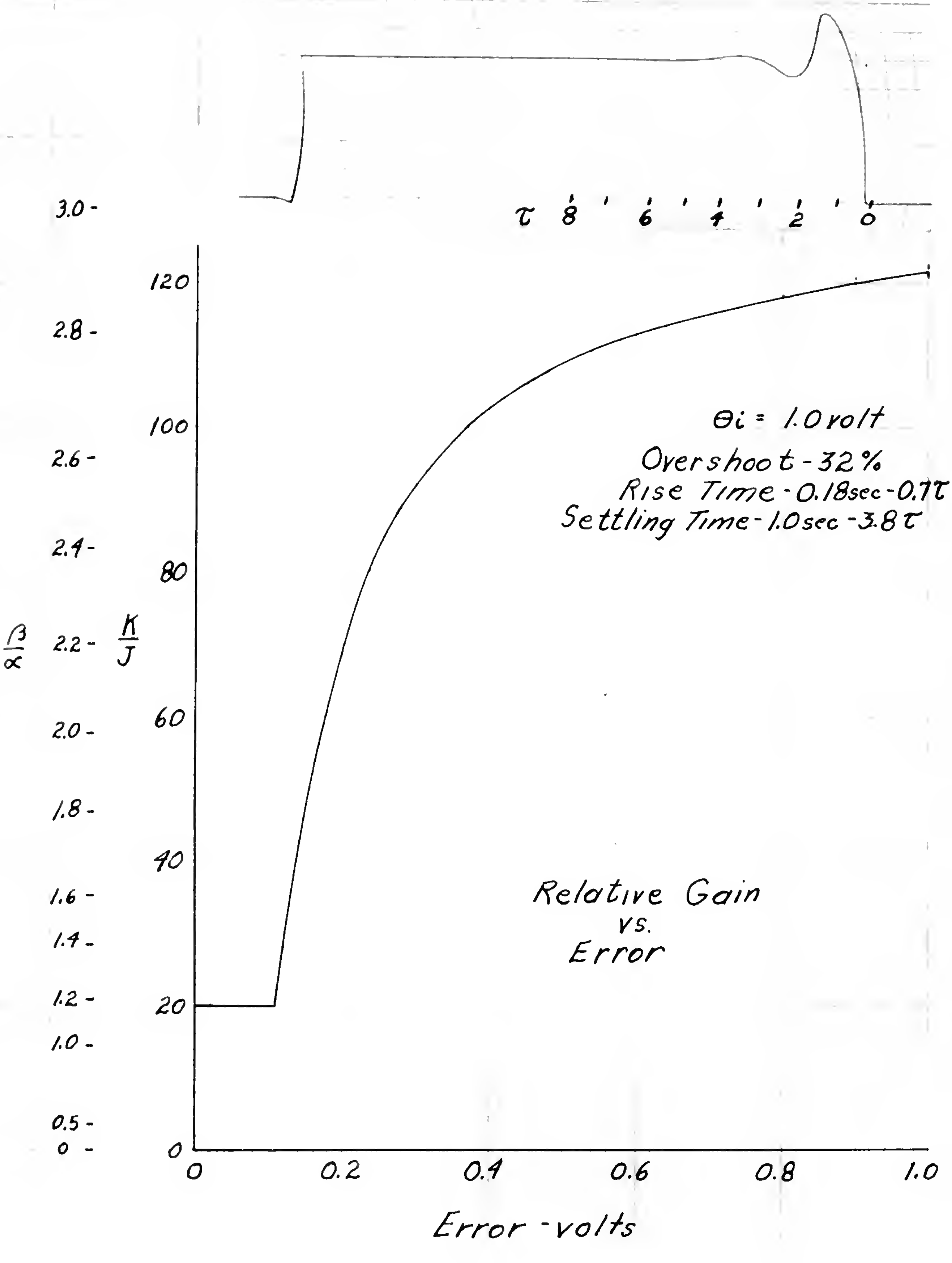


Figure XXXII

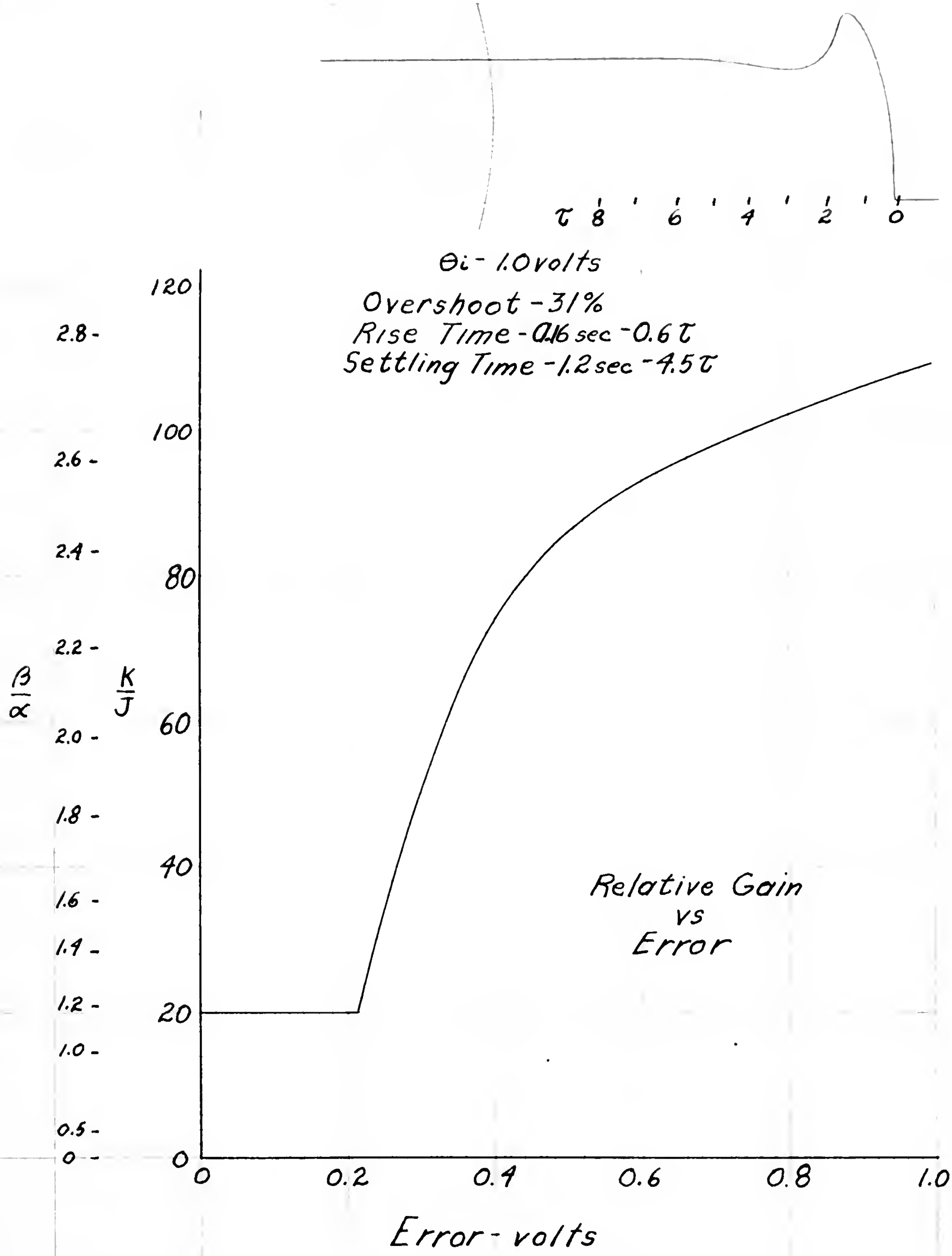


Figure XXXIII

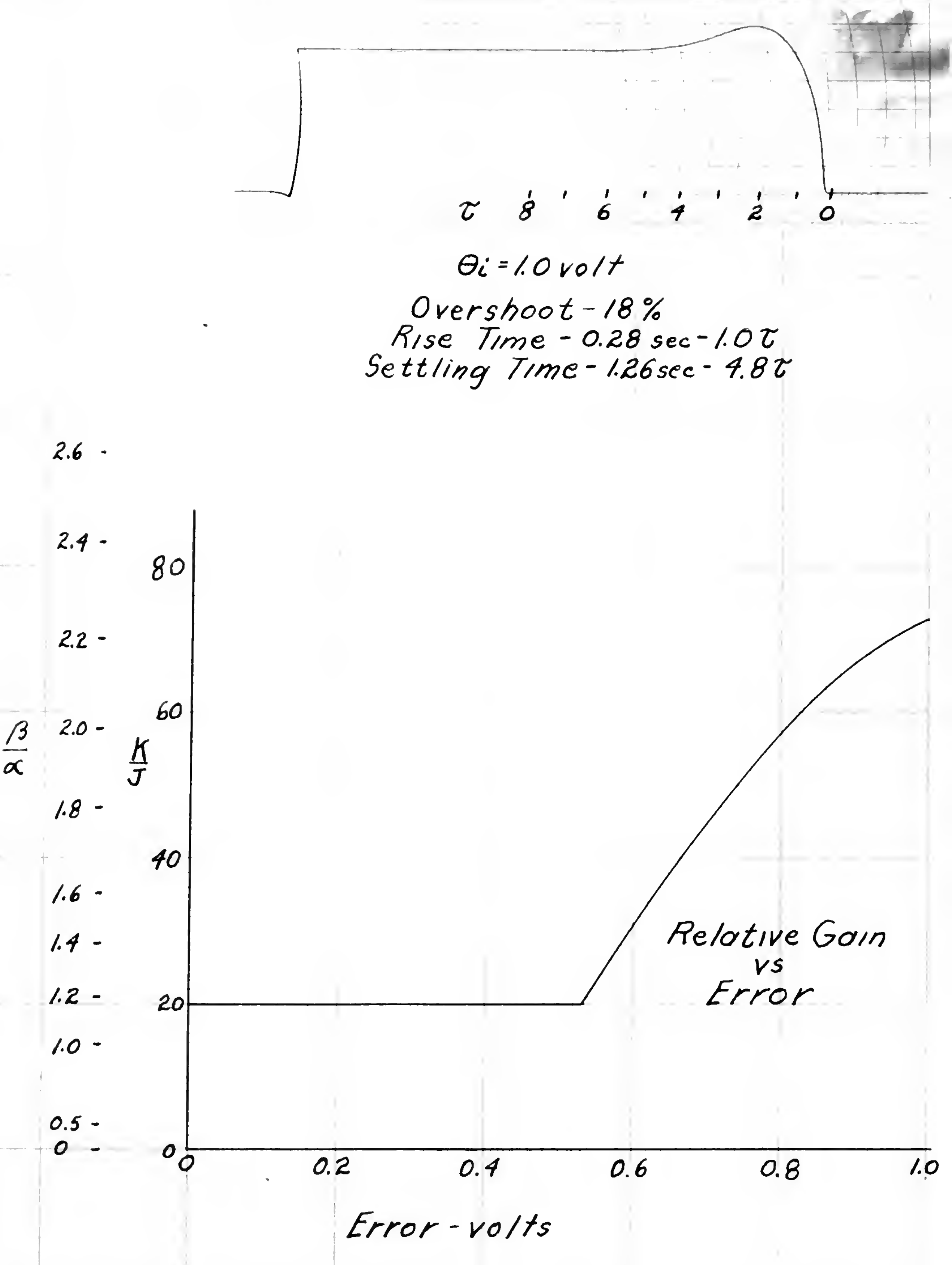


Figure XXXIV

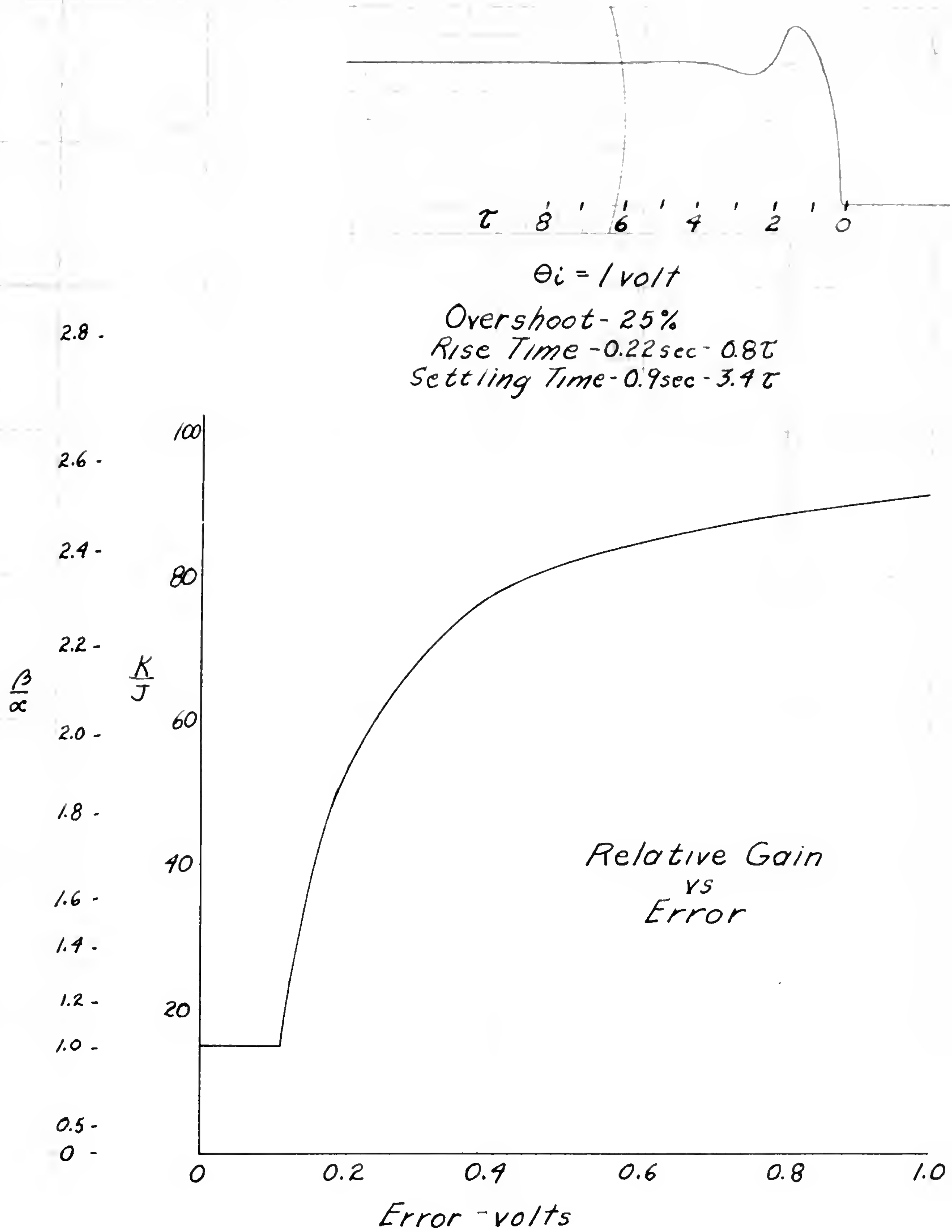

τ 8 6 4 2 0
θi = 1 volt
Overshoot - 25%
Rise Time - 0.22 sec - 0.8τ
Settling Time - 0.9 sec - 3.4τ
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.5
0
β/α
K/J
100
80
60
40
20
Relative Gain vs Error
0
0.2
0.4
0.6
0.8
1.0
Error - volts

Figure XXXV

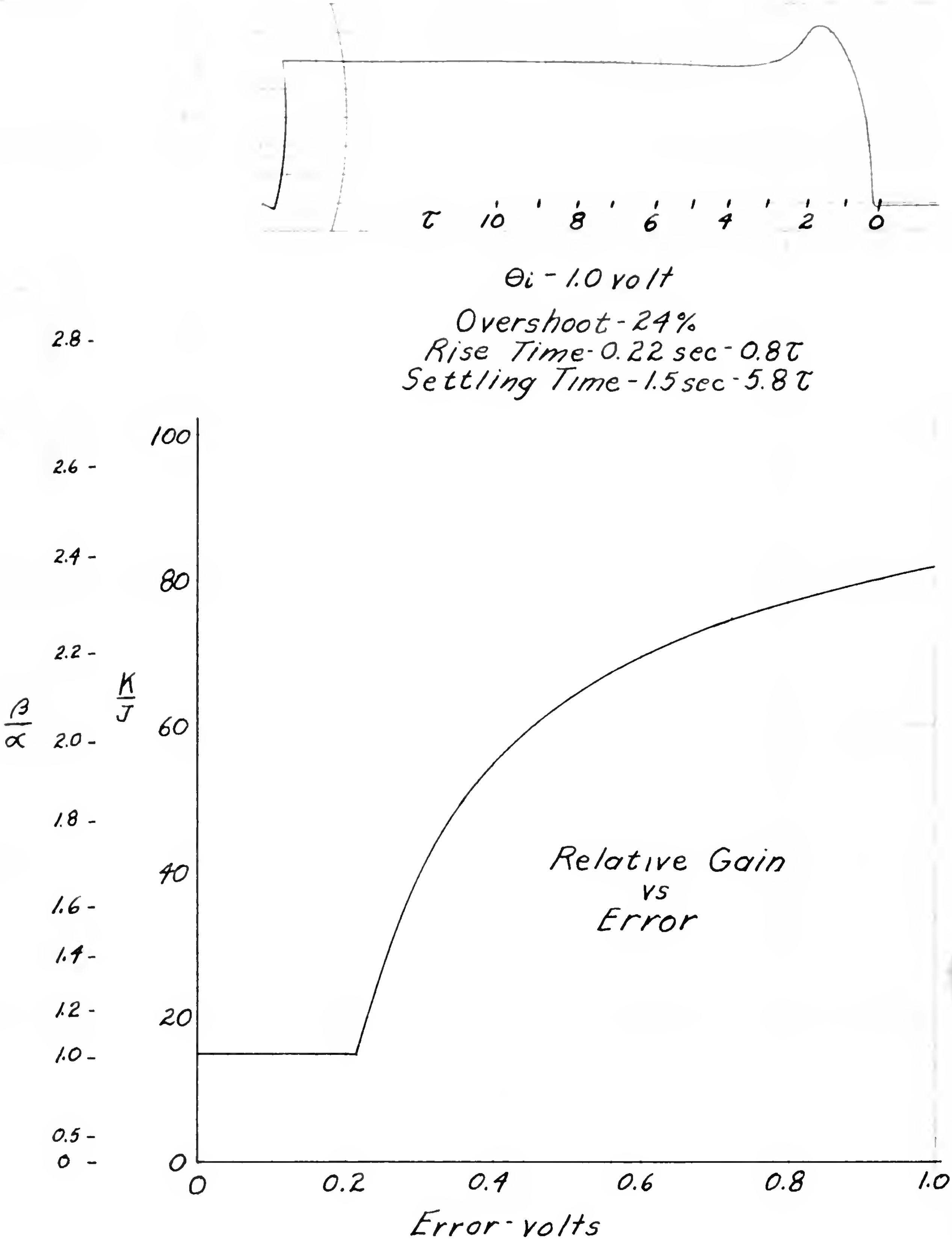
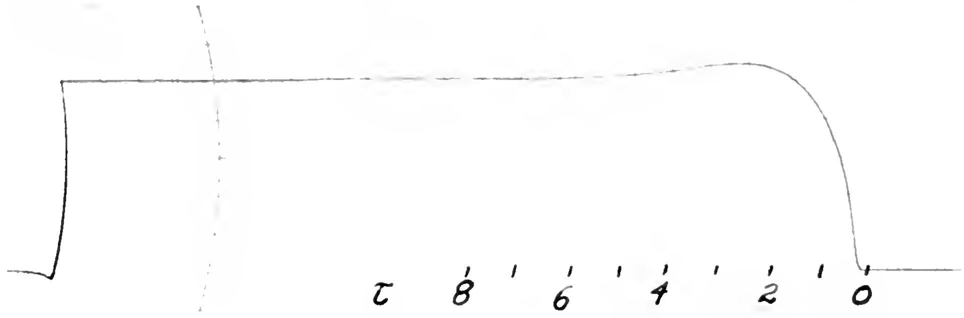


Figure XXXVI



θ_i - 1.0 volts

Overshoot - 9 %

Rise Time - 0.40 sec - 1.5 τ

Settling Time - 1.3 sec - 4.9 τ

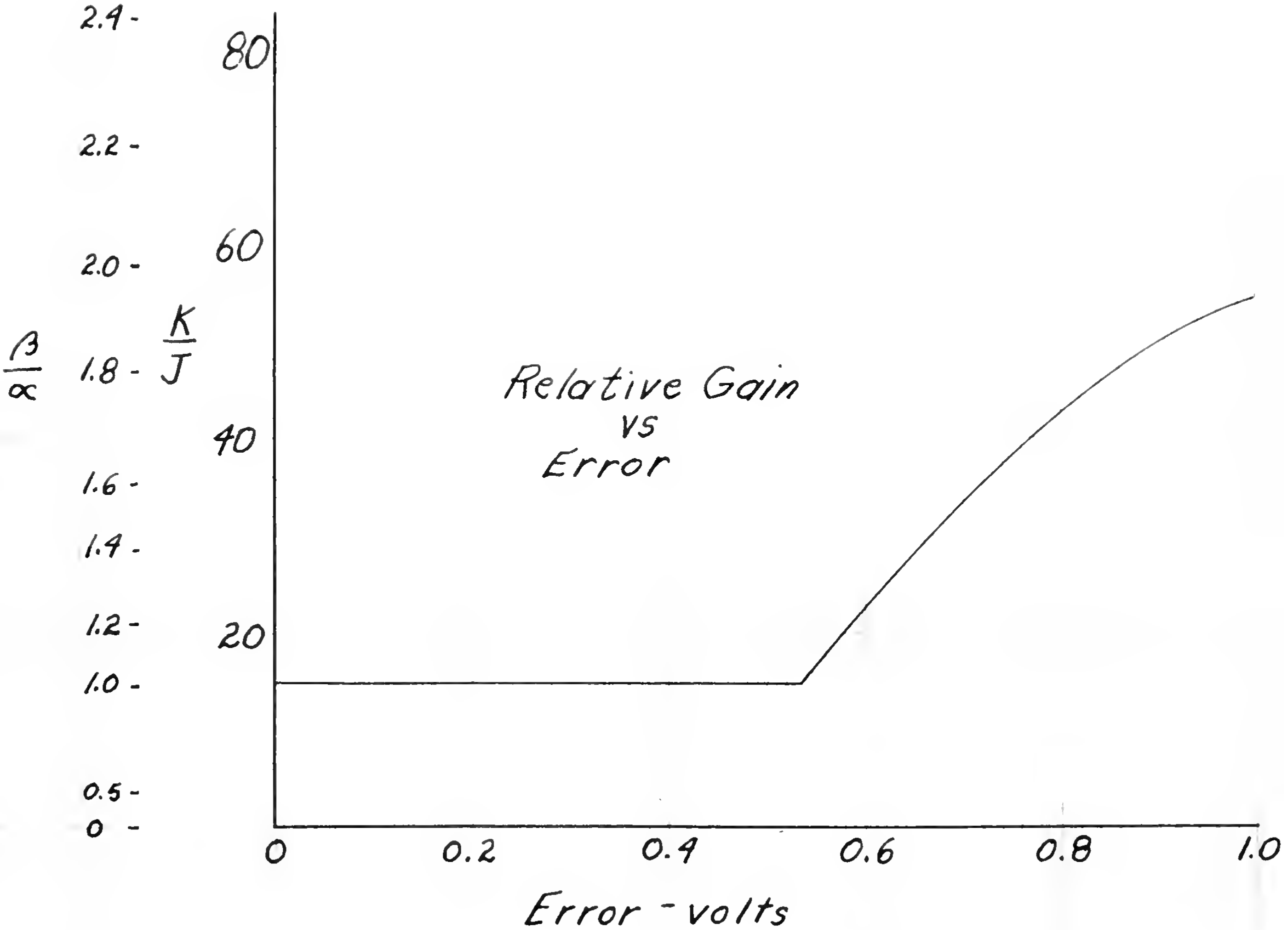
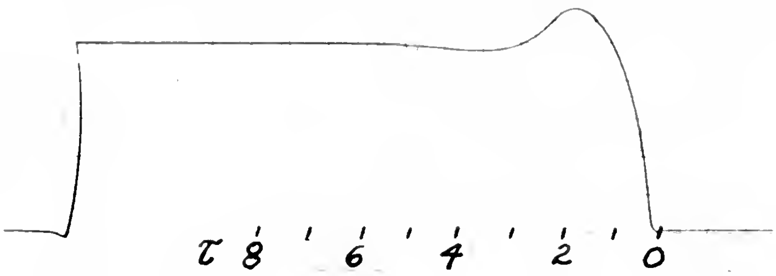


Figure XXXVII



θ_i = 1.0 volts
Overshoot - 18 %
Rise Time - 0.28 sec - 1.1 τ
Settling Time - 1.2 sec - 4.5 τ

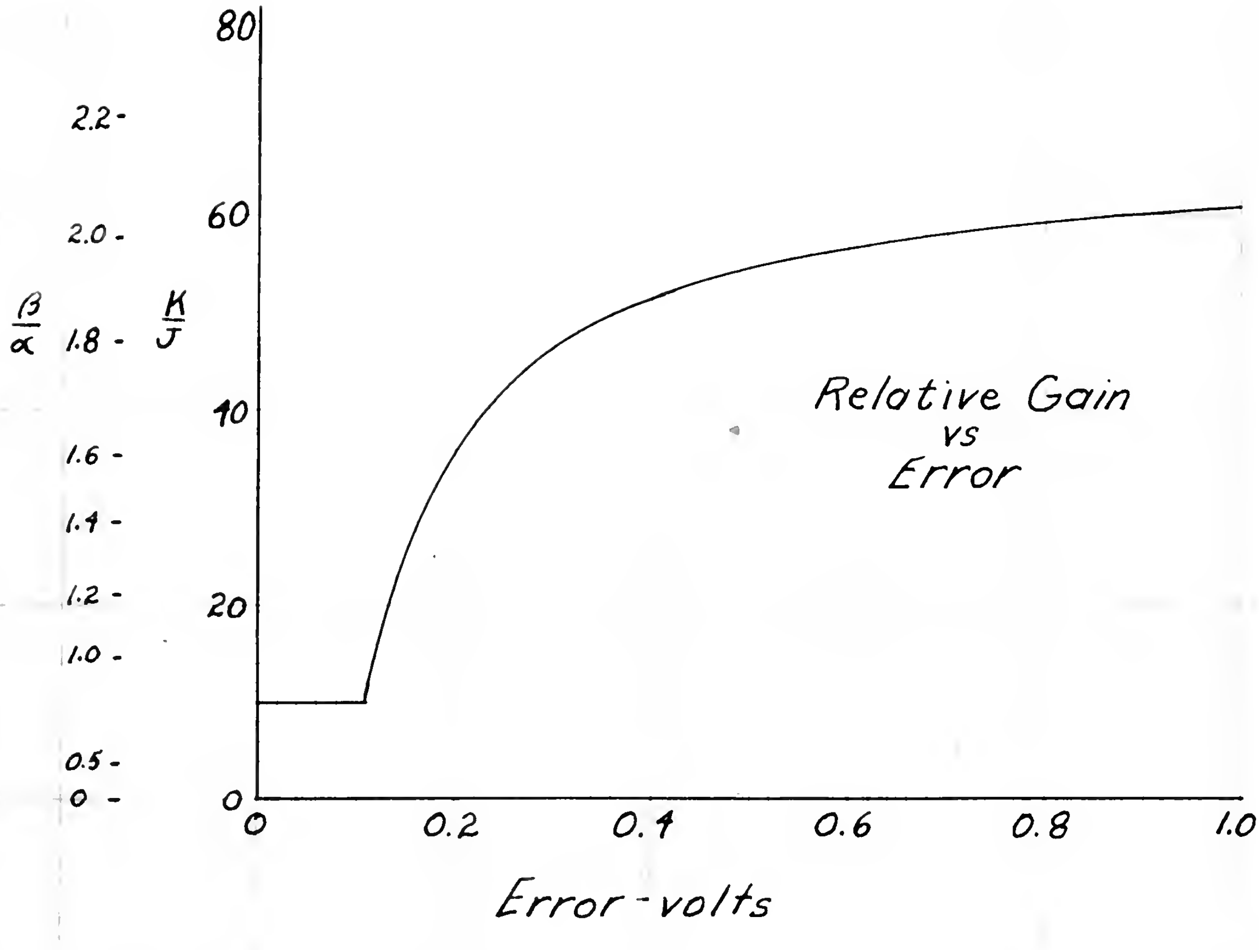
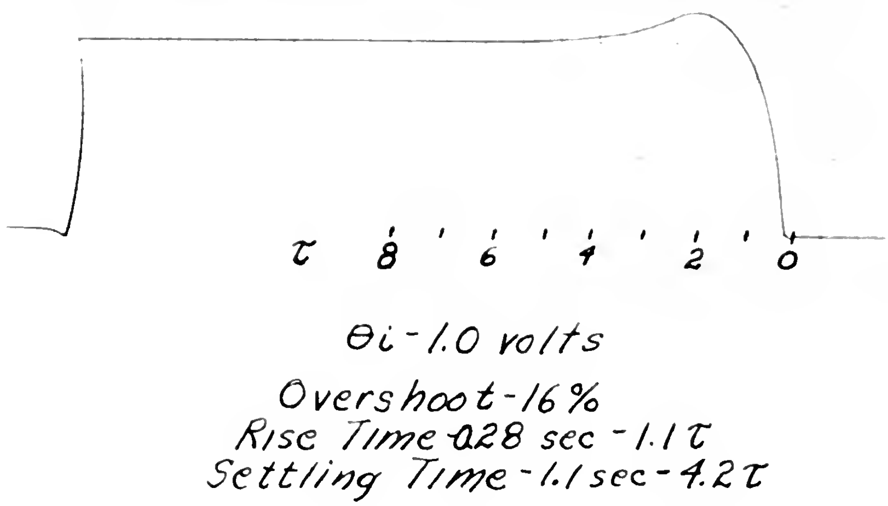
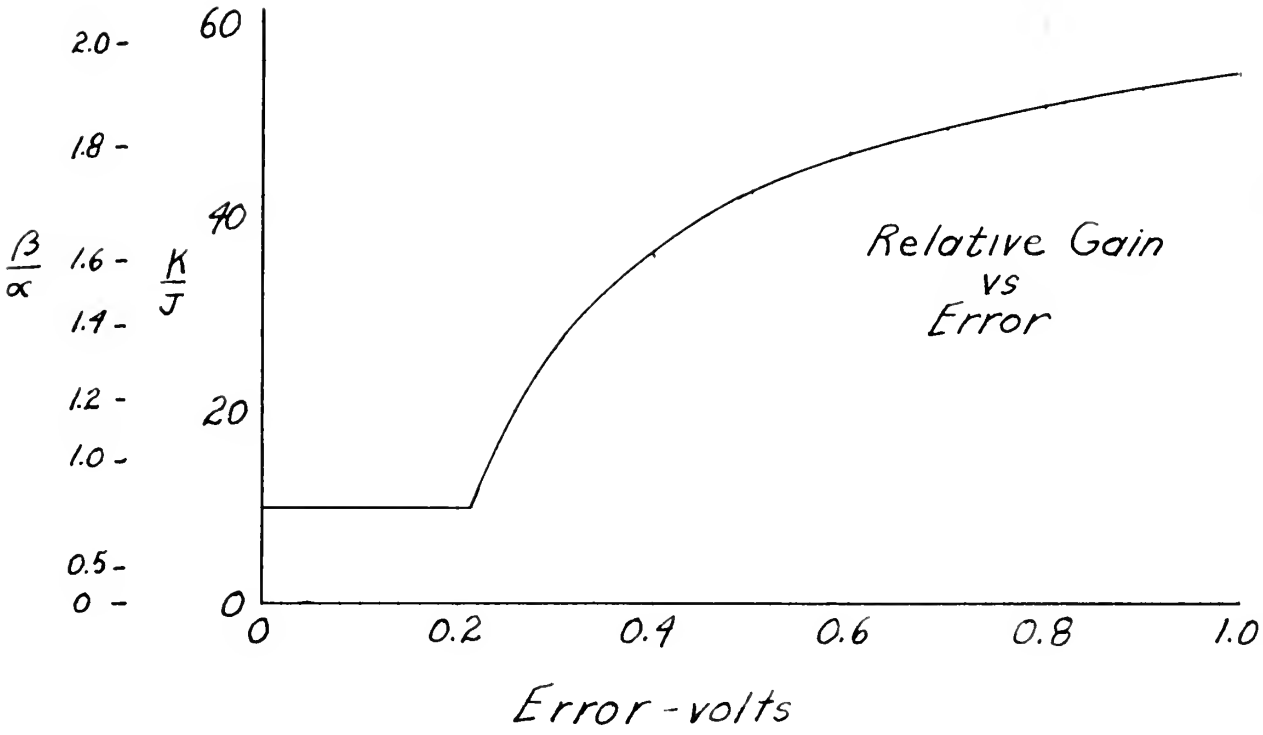


Figure XXXVIII

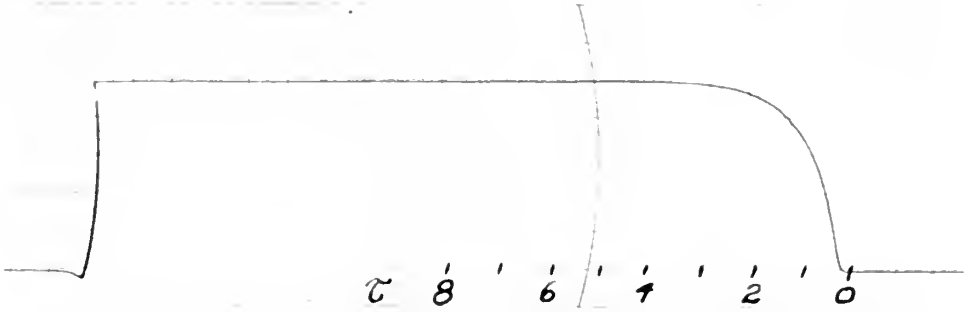


τ 8 6 4 2 0
θi - 1.0 volts
Overshoot - 16%
Rise Time 0.28 sec - 1.1τ
Settling Time - 1.1 sec - 4.2τ



β/α
2.0
1.8
1.6
1.4
1.2
1.0
0.5
0
K/J
60
40
20
0
0
0.2
0.4
0.6
0.8
1.0
Relative Gain vs Error
Error - volts

Figure IXL



θ_i = 1.0 volts

Overshoot - none

Settling Time - 0.8 sec - 3.0τ

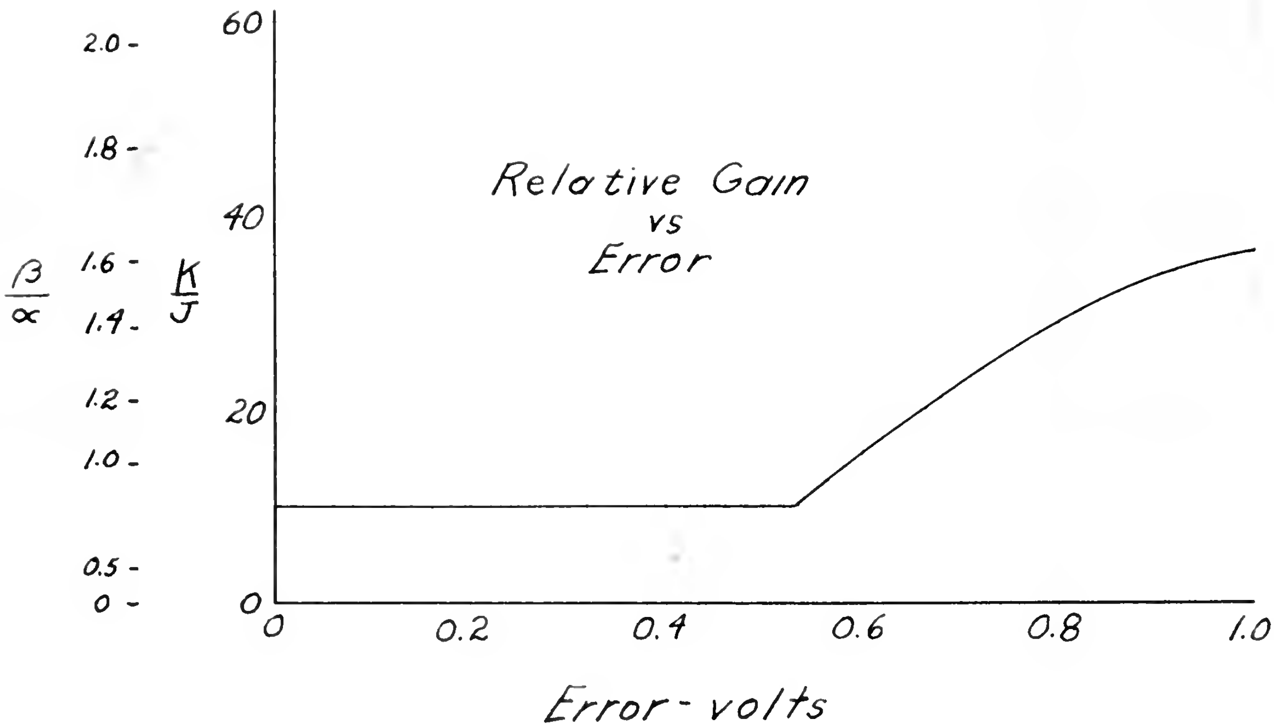


Figure XL

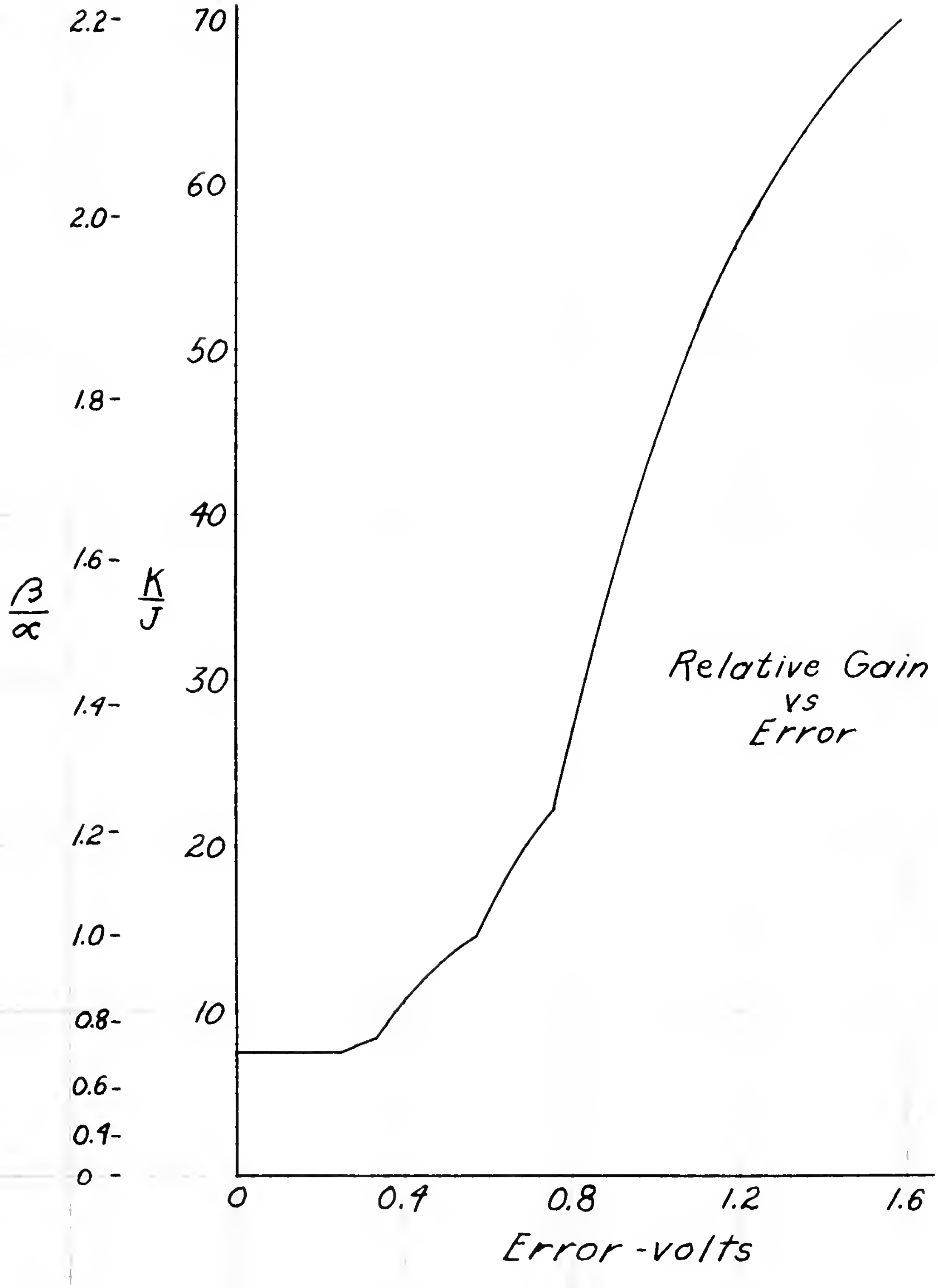


Figure $\overline{\underline{XLI}}$-a

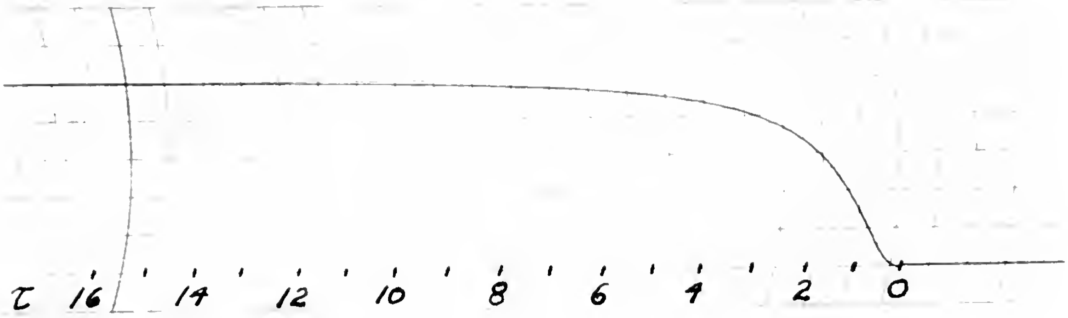


Figure XLI-b

θ_i - 0.5 volts
overshoot - none
Settling time - 2.5 sec - 9.5 τ

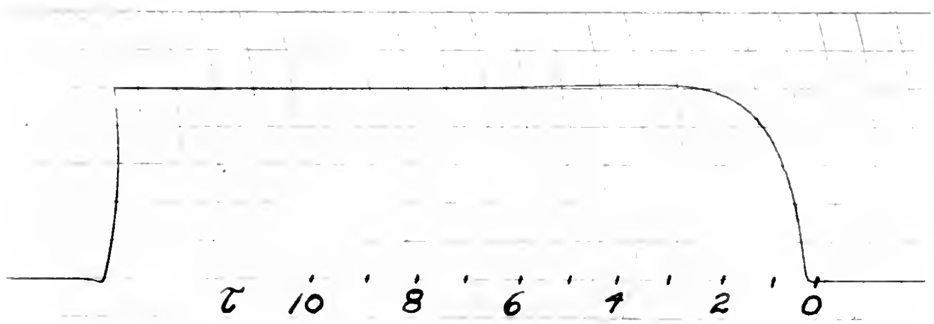


Figure XLI-c

θ_i - 1.0 volts
Overshoot - 2 %
Rise time - 0.6 sec. - 2.3 τ
Settling time - 0.58 sec - 2.2 τ

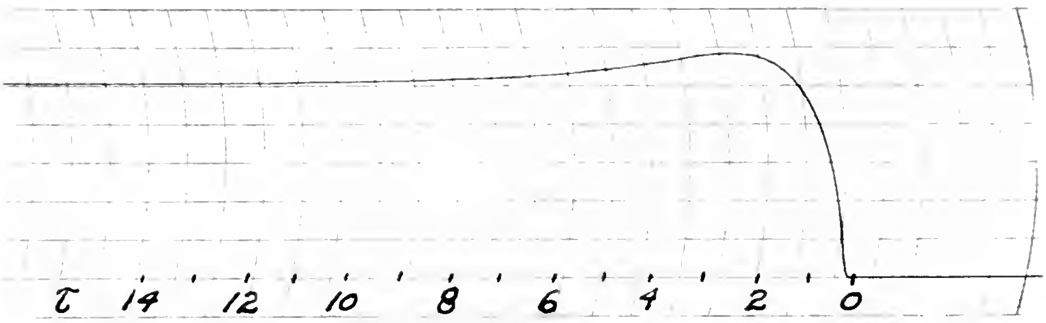


Figure XLI-d

θ_i - 1.5 volts
Overshoot - 17 %
Rise time - 0.3 sec - 1.1 τ
Settling time - 2.5 sec - 9.5 τ

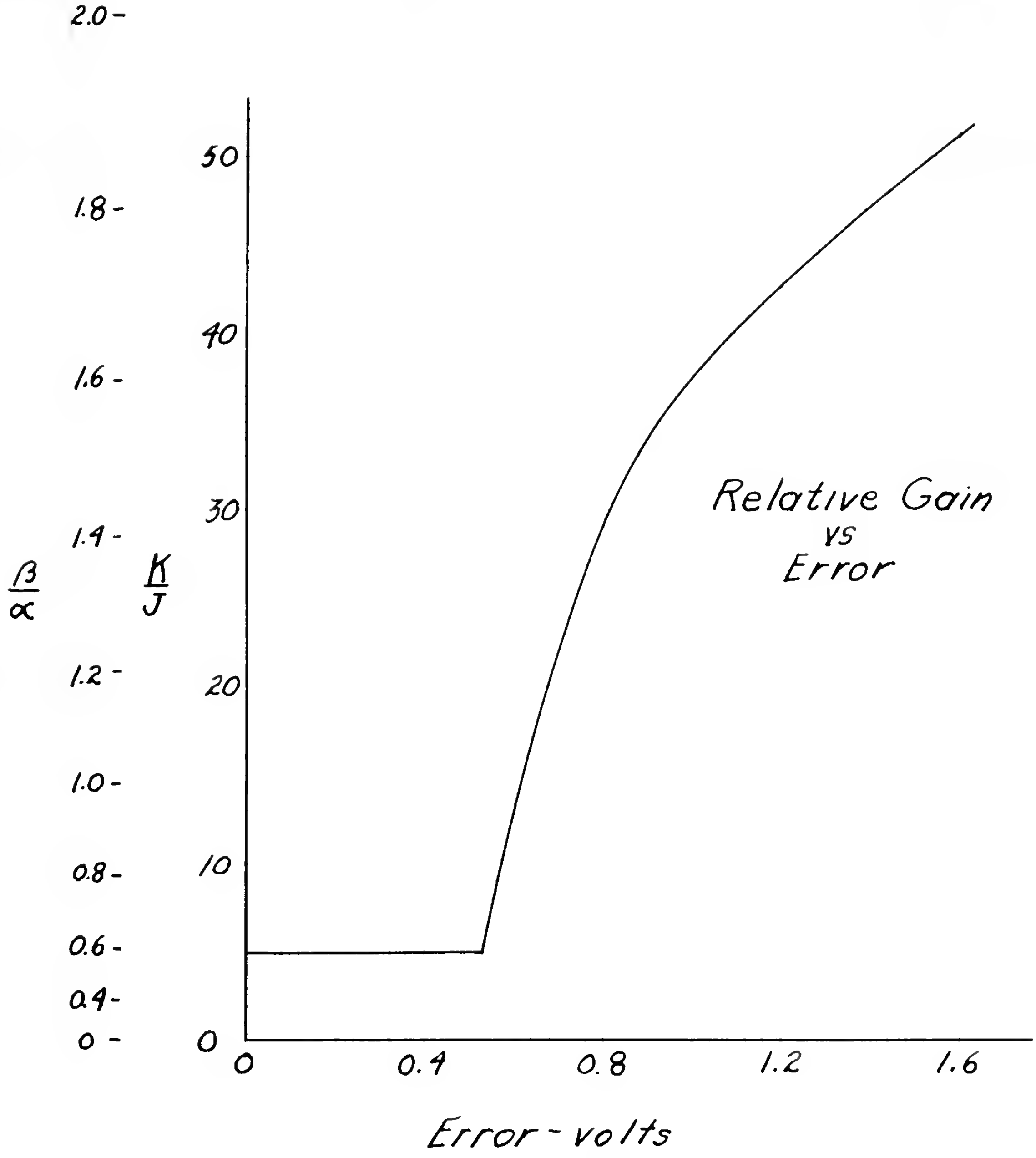


Figure XLII-a

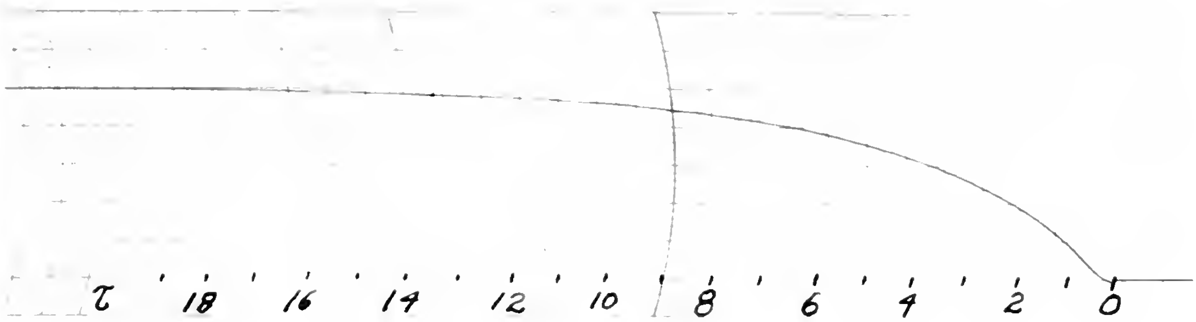


Figure XLII-b

θ_i - 0.5 volts
Overshoot - none
Settling Time - 4.2 sec. - 16 τ

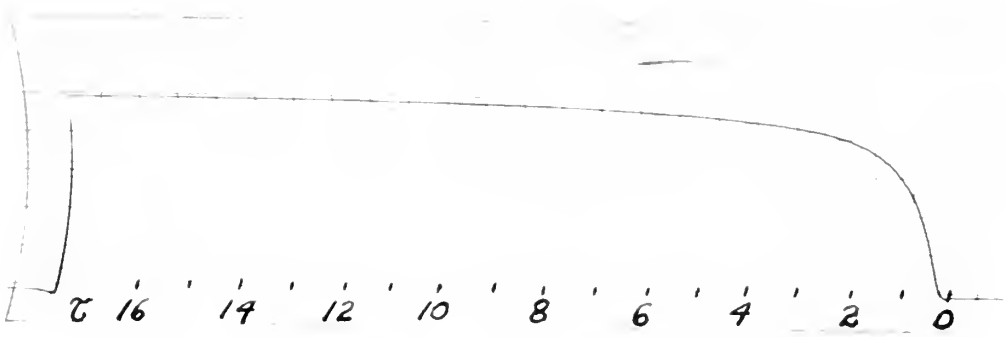


Figure XLII-c

θ_i - 0.75 volts
overshoot - none
Settling Time - 3.2 sec. - 12 τ

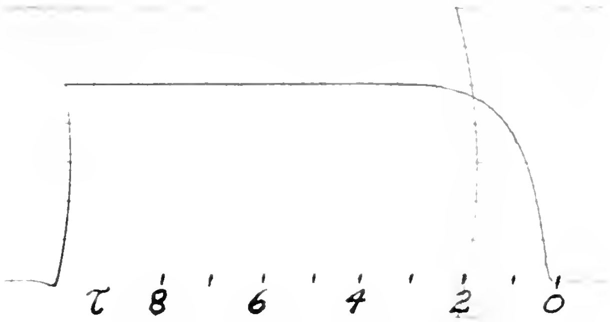


Figure XLII-d

e_i - 1.0 volts
Overshoot - nil
Settling Time - 0.58 sec - 2.2τ

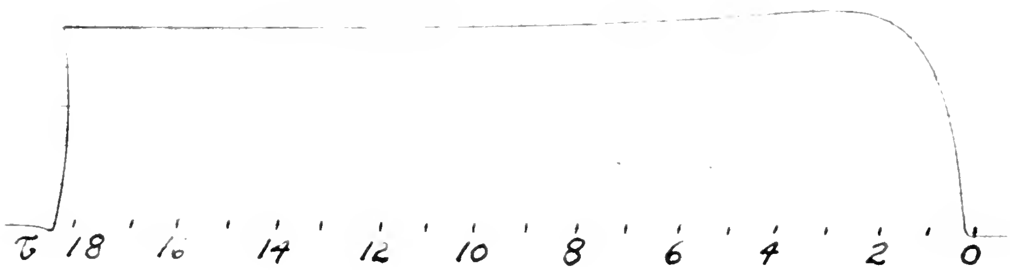


Figure XLII-e

e_i - 1.0 volts
Overshoot - 13 %
Rise Time - [illegible] sec. - 1.4 τ
Settling Time - 3.0 sec. - 11.4τ